Function Space

Tan-Jianbin

School of Mathematics
Sun Yat-sen University

Seminar on Statistics 105c

Function Space

1. Continuous function and L^p Space
   - "Representation" type Theorem
   - Stone-Weierstrass Theorem
   - Integral Operator and Mercer's Theorem
2. Reproducing Kernel Hilbert Space
   - Kernel Trick and Moore-Aronszajn Theorem
   - Integral Representation Theory
   - Sum and product of RKHSs
   - Operators on RKHSs
3. Sobelev Space

Definition

K is a compact metric space, $C(K)$ is a function space that consists of all the continuous functions $f : K \to R$.

We consider K are some closed intervals $[a, b]$, without loss of generality, we focus on $[0, 1]$.

Theorem

$C([0, 1])$ is a normed vector space, equipped with a norm $|| \cdot ||_p$: $||f||_p = (\int |f(x)|^p dx)^{1/p}$

One can show that $p = \infty$, $||f||_p = \sup\{|f(x)|\}$, we mark that $|| \cdot ||_{sup}$.

Theorem

$(C([0,1]), ||\cdot||_{sup})$ *is a Banach space.*

Proof.

Take a Cauchy $\{f_n\}$. Then $\forall x \in [0,1]$, $\{f_n(x)\}$ Cauchy since $\sup_{n,m\geq N}|f_n(x) - f_m(x)| \to 0 \Rightarrow y_x \in R$, $f_n(x) \to y_x$.

Let $f : [0,1] \to R$, $f(x) = y_x$. f is continuous at $[0,1]$ since, $|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)|$

And $f_n \to f$ because $\forall \varepsilon > 0$, $\exists\ N > 0$, $n, m \geq N$, $\forall x \in [0,1]$ $|f_m(x) - f_n(x)| \leq \varepsilon \Rightarrow sup_{x\in[0,1]}|f(x) - f_n(x)| \leq \varepsilon$ □

Theorem

$A \subset\subset C(K) \Leftrightarrow$ *A is bounded and equi-continuous.*

Proof.

"$\Rightarrow$": A is a totally bound set $\Rightarrow \forall \varepsilon > 0$, $\exists$ finite $\{f_n\} \subset A$, $\forall f \in A$, $\exists i$, $||f - f_i||_{sup} \leq \varepsilon/3$, then A is equi-continuous because $|f(x) - f(y)| \leq |f(x) - f_i(x)| + |f_i(x) - f_i(y)| + |f_i(y) - f(y)|$

"$\Leftarrow$": K totally bounded $\Rightarrow \forall \varepsilon > 0$ $\exists$ finite $\{x_n\}$, $\forall x \in K$, $\exists i$ $d(x, x_i) \leq \varepsilon$. Then $\forall$ distinct $\{f_n\} \subset A$, $\because$ $\{f_n(x_1)\}$ is bounded,

$\exists$ $\{f_n^{(1)}(x_1)\}$ converges $\Rightarrow$ $\{f_n^{(1)}\}$ converges at x_1 $\Rightarrow$ $\exists\{f_n^{(2)}\} \subset \{f_n^{(1)}\}$ converges at $x_2 \Rightarrow \exists$ $\{f_n^{(k)}\}$ converges at $\{x_n\}_{n=1}^k$. $\{f_n^{(k)}\}$ Cauchy since $|f_n^{(k)}(x) - f_m^{(k)}(x)| \leq |f_n^{(k)}(x) - f_n^{(k)}(x_i)| + |f_n^{(k)}(x_i) - f_m^{(k)}(x_i)| + |f_m^{(k)}(x_i) - f_m^{(k)}(x)|$ □

Definition

$(E, \mathcal{F}, \mu)$ is a measure space. $1 \leq p \leq \infty$. L^p is a function space that consists of all the measurable functions $f : E \to R$ s.t. $(\int |f|^p d\mu)^{1/p} < \infty$. Define $||f||_p = (\int_A |f|^p d\mu)^{1/p}$ and $||f||_\infty = ess\sup\{f(x); x \in E\} = inf_{\mu(A)=0} sup\{f(x); x \in E - A\}$.

If E compact metric space. $\forall f \in C(E)$, $||f||_{sup} = ess\sup\{f\}$ since that $f(E)$ is compact, then $\{x; f(x) > M\}$ is non-empty open set for $M < ||f||_{sup} \Rightarrow \mu(\{x; f(x) > M\}) > 0$.

Theorem

$1 \le p, q \le \infty$, $f_1 \in L^p$, $f_2 \in L^q$ *with* $1/p + 1/q = 1$. *Then* $||f_1 f_2||_1 \le ||f_1||_p \, ||f_2||_q$

Proof.

If $p = \infty, q = 1$, $\int |f_1 f_2| d\mu \le ||f_1||_\infty \int |f_2| d\mu$
For $1 \le p, q < \infty$, we have $a_1 a_2 \le p^{-1} a_1^p + q^{-1} a_2^q$
Let $a_1 = f_1(x)/||f_1||_p$, $a_2 = f_2(x)/||f_2||_q$ and integrate. □

Theorem

$1 \le p \le \infty$, $f, g \in L^p$. *Then* $||f+g||_p \le ||f||_p + ||g||_p$

Proof.

$$||f+g||_p^p = \int |f+g|^p d\mu = \int |f+g|^{p-1}|f+g| d\mu$$
$$\le \int |f+g|^{p-1}|f| d\mu + \int |f+g|^{p-1}|g| d\mu$$
$$\le ||f+g||_p^{p/q}||f||_p + ||f+g||_p^{p/q}||g||_p$$

□

Corollary

$(L^p, ||\cdot||_p)$ *is normed vector space, for* $1 \le p \le \infty$.

Theorem

$(L^p, ||\cdot||_p)$ is a Banach space, for $1 \leq p \leq \infty$.

Proof.

Take Cauchy $\{f_n\} \subset L^p$, $sup_{n,m \geq N}||f_m - f_n||_p \to 0$. Then $\exists n_k > 0$, $||f_{n_{k+1}} - f_{n_k}||_p \leq 1/2^k \Rightarrow \sum_k ||f_{n_{k+1}} - f_{n_k}||_p < \infty \Rightarrow$ $||\sum_k (f_{n_{k+1}} - f_{n_k})||_p < \infty \Rightarrow \sum_k |f_{n_{k+1}}(x) - f_{n_k}(x)| < \infty$ a.s. $\Rightarrow \{f_{n_k}(x)\}$ Cauchy a.s. Let $f(x) = \lim_k f_{n_k}(x)$ a.s.

$||f - f_n||_p = ||\liminf_k \ f_{n_k} - f_n||_p \leq \liminf_k ||f_{n_k} - f_n||_p$, $\forall \varepsilon > 0$, if n_k, n big enough, $||f_{n_k} - f_n||_p \leq \varepsilon$. $\square$

Theorem

$\mu(E) < \infty$, $1 \leq p_1 \leq p_2 < \infty$, then $L^\infty \subseteq L^{p_2} \subseteq L^{p_1}$
f is a measurable function, then $||f||_p \to ||f||_\infty$
If $\mu(E) = 1$, then $||f||_p \uparrow ||f||_\infty$

Proof.

$$\int |f|^{p_1} d\mu \leq ||f||_{p_2}^{p_1} \mu(E)^{1-p_1/p_2} \overset{\mu(E)=1}{\Rightarrow} ||f||_{p_1} \leq ||f||_{p_2}$$
$$(\int |f|^{p_1} d\mu)^{1/p_1} \leq ||f||_\infty (\mu(E))^{1/p_1} \overset{\mu(E)=1}{\Rightarrow} \lim_{p_1} ||f||_{p_1} \leq ||f||_\infty$$
$$\forall \varepsilon > 0, \text{ let } A_\varepsilon = \{f(x) \geq ||f||_\infty - \varepsilon\} \Rightarrow \mu(A_\varepsilon) > 0$$
$$||f||_{p_1} \geq (\int_{A_\varepsilon} |f|^{p_1} d\mu)^{1/p_1} \geq (||f||_\infty - \varepsilon)\ (\mu(A_\varepsilon))^{1/p_1}$$
$$\text{Then } ||f||_\infty - \varepsilon \leq \lim_{p_1} ||f||_{p_1} \Rightarrow ||f||_\infty \leq \lim_{p_1} ||f||_{p_1}$$
□

Theorem

$C([0,1])$ is dense in $L^p([0,1])$, for $1 \leq p < \infty$.

Proof.

$\forall A \in \mathcal{B}([0,1])$, for $\forall \varepsilon > 0$, $\exists$ a compact set $B \subset A$ s.t. $\mu(B - A) < \varepsilon/2 \Rightarrow ||I_A - I_B||_p < \varepsilon/2$.

Let $f_n(x) = 1/(d(B,x)+1)^n$, then $f_n \to I_B$ and $f_n \in C([0,1])$ $\Rightarrow ||I_A - f_n||_p < \varepsilon$. $\square$

Definition

$BV[0,1]$ is a function space on $[0,1]$ that $\forall f \in BV[0,1]$, $V(f) = \sup \sum_i |f(t_i) - f(t_{i-1})| < \infty$.

If $f^{(1)} \in C([0,1])$, then $V(f) = \int |f^{(1)}(x)|dx$.

$V(f|_{[0,x]}) \pm f(x)$ is a non-decreasing function since that $V(f|_{[0,x]}) - V(f|_{[0,y]}) \geq V(f|_{[y,x]}) \geq |f(x) - f(y)|$.

Theorem

$f \in BV[0,1] \Leftrightarrow \exists$ two finite non-decreasing functions f_1, f_2, $f = f_1 - f_2$.

Theorem

Define a norm on $BV[0,1]$: $||f||_{BV} = V(f) + |f(0)|$, then $(BV[0,1], ||\cdot||_{BV}))$ is Banach space.

Proof.

Since $|f_n(x) - f_m(x)| \leq ||(f_n - f_m)|_{[0,x]}||_{BV} \leq ||f_n - f_m||_{BV}$, we know that if $\{f_n\}$ Cauchy then $\{f_n(x)\}$ is Cauchy, so we can define f, $f(x) = \lim_n f_n(x) \Rightarrow f \in BV[0,1]$.

$\sum_i |f(t_i) - f(t_{i-1}) - f_n(t_i) + f_n(t_{i-1})| \to 0 \ \forall \ t_i \in [0,1] \Rightarrow ||f - f_n||_{BV} \to 0.$ □

Theorem

T is a bounded linear functional on $C[0,1] \Leftrightarrow \exists\ \mu \in BV[0,1]$, $Tf = \int_0^1 f\ d\mu$, $\forall f \in C[0,1]$.

Proof.

"$\Rightarrow$": $B[0,1]$: bound functions on $[0,1]$, $C[0,1] \subset B[0,1]$. $\exists\ \hat{T}$, which domain is $B[0,1]$ and $||\hat{T}|| = ||T||$.

Define $f_n \in B[0,1]$, $|t_n - t_{n-1}| \leq 1/n$, $f_n(t) = \sum_n f(t_{n-1})$ $I_{[t_{n-1},t_n)}(t)$, $||f_n - f|| \to 0$.

$\hat{T} f_n = \sum_n f(t_{n-1}) \hat{T} I_{[t_{n-1},t_n)} := \sum_n f(t_{n-1})(\mu(t_n) - \mu(t_{n-1}))$ $\Rightarrow Tf = \int f\ d\mu$, if $\mu \in BV[0,1]$.

$\sum_n |\mu(t_n) - \mu(t_{n-1})| = \sum_n |\hat{T} I_{[t_{n-1},t_n)}| = \sum_n sign(\hat{T} I_{[t_{n-1},t_n)})$ $\hat{T} I_{[t_{n-1},t_n)} \leq ||\hat{T}||\ ||\sum_n sign(\hat{T} I_{[t_{n-1},t_n)}) I_{[t_{n-1},t_n)}|| \leq ||\hat{T}||$. □

Define $Tg = \int fg \ d\mu$, $f \in L^q$, $\forall g \in L^p$.

Note that this is a bounded functional of L^p since $|\int fg \ d\mu| \leq ||f||_q \ ||g||_p$ and $||T|| \leq ||f||_q$. Let $g = |f|^{q-1} sign(f)/||f||_q^{q-1} \Rightarrow ||T|| = ||f_q||$.

Moreover, one can show that $(L^p)^* \approxeq L^q$. Roughly speaking, $\forall \ T \in (L^q)^*$, $\nu(A) = TI_A \Rightarrow Tf = \int f \ d\nu = \int fg \ d\mu$.

Define a inner product of L^2, $\langle f_1, f_2 \rangle_2 = \int f_1 f_2 \, d\mu$, L^2 is a Hilbert space. The property of $L^2([0,1])$ depends on $C([0,1])$.

Definition

Define a Bernstein operator $B_n : C([0,1]) \to R_n[x]$, $B_n(f) = Ef(Y_x/n) = \sum_k f(k/n) C_n^k \, x^k (1-x)^{n-k}$, $Y_x \sim B(n,x)$.

$P_n[x]$ consists of all the polynomials which degree $\leq n$.

Now we use Bernstein operator to deduce that $R[x]$ is a dense subset of $C([0,1])$, $\forall 1 \leq p \leq \infty$.

Theorem

$P[x]$ is a dense subset of $C([0,1])$.

Proof.

$\forall \varepsilon, \delta > 0$, let $A^\delta = \{|f(Y_x/n) - f(x)| \geqslant \delta\}$. Then $|(B_n f)(x) - f(x)| \leq E|f(Y_x/n) - f(x)| \leq E(I_{A_\delta}|f(Y_x/n) - f(x)|) + \delta$. $\exists \delta_1$ s.t. $P(A^\delta) \leq P(|Y_x/n - x| \geq \delta_1) \leq (1-x)x(n\delta_1^2)^{-1} \leq (4n\delta_1^2)^{-1}$ $\Rightarrow E(I_{A_\delta}|f(Y_x/n) - f(x)|) \leq 2M(4n\delta_1^2)^{-1}$. □

Use G-S to $\{1, x, ...\}$, we get the Legendre polynomials, which is the COB of $L^2([0,1])$.

Corollary

$\{f_0(x) = 1, f_n(x) = \sqrt{2}\cos(n\pi x)\}$ and $\{\sqrt{2}\sin(n\pi x)\}$ is a COB of $L^2([0,1])$.

Proof.

$g \in C([0,1])$, let $k(\cos\pi x) = g(x) \Rightarrow k(x) = g(cos^{-1}(x)/\pi)$ $\Rightarrow k \in C([0,1])$. Then $\exists$ polynomial p s.t. $||p-k||_{sup} \leq \varepsilon \Rightarrow$ $|p(\cos\pi x) - k(\cos\pi x)| \leq \varepsilon$, $\forall x \in [-1,1]$. Let $h(x) = p(\cos\pi x)$, $||h-g||_{sup} \leq \varepsilon$.

Let $h(x) = g(x)(\sin\pi x)^{-1}$. $\forall \varepsilon > 0$, $\exists k(x) = \sum_n a_n \cos(n\pi x)$, $||h-k||_2 \leq \varepsilon/2$. And $||h-k||_2^2 = \int (f(x)/\sin\pi x - k(x))^2 dx$ $\geq \int (f(x) - k(x)\sin\pi x)^2 dx$. We can define $g_\delta = gI_{(\delta,1]}$ instead of g to avoid the invalidity of $h(0)$. □

Theorem

If $\Lambda \subset C[K]$ is a unital sub-algebra, then Λ is dense in $C(K)$ if and only if $\forall$ distinct $x, y \in K$, $\exists f \in \Lambda$ s.t. $f(x) \neq f(y)$.

Let $E = \{e^{2\pi ix} : 0 \leq x \leq 1\}$. One can think of this as $[0, 1]$ with 0 and 1 identified. Then $\{1, \sqrt{2}\sin(2n\pi x), \sqrt{2}\cos(2n\pi x)\}$ are COB for $L^2([0, 1])$.

E compact set, define $\Lambda = \{\sum_{n=1}^{k} f_n(s)g_n(t); g, f \in C(E)\}$. We call functions in Λ degenerate kernels and Λ is dense in $C(E \times E)$.

Definition

K is a measurable kernel on $(E \times E, \mathcal{F} \times \mathcal{F}, \mu \times \mu)$ s.t. $C = \int\int_{E\times E} K^2(s,t)d\mu(s)d\mu(t) < \infty$.

Define $T : L^2 \to L^2$, $T(f) = \int_E K(s,t)f(t)d\mu(t)$.

Assumed that E is compact, we focus on integral operators T which kernels K are continuous on $E \times E$ and $\mu(E) < \infty$.

Property

T is bounded and $T(L^2(E)) \subset C(E)$.

Proof.

$|Tf(s)| = |\int_E K(s,t)f(t)d\mu(t)| \leq ||K(s,\cdot)||_2||f||_2 \Rightarrow$ $||Tf||_2 \leq C||f||_2 \Rightarrow T \in B(L^2(E))$. □

Moreover, $T(L^2(E)) \subset\subset C(E) \Rightarrow T \in K(L^2(E), C(E))$. And $f \in C([0,1])$, $\sup_{||f||_{sup}=1} |G_t(f)| = |f(t)| \leq 1$.

Definition

A kernel K is a bivariate function $K : E \times E \to R$
Symmetric: $K(s,t) = K(t,s)$, $s, t \in E$
Non-negative definite: $\forall$ finite $\{a_n\} \in R$ and $\{t_n\} \in E$,
$\sum_i \sum_j a_i a_j K(t_i, t_j) \geq 0$

If the kernel of $T \in B(L^2(E))$ exist, K is unique since for other H s.t. $\int (K(s,t) - H(s,t)) f(s) d\mu(s) = 0$, $\forall f \in L^2(E)$, $K = H$.

Property

K is symmetric $\Leftrightarrow$ T is self-adjoint.

Proof.

"$\Rightarrow$": $\langle Tf, g\rangle_2 = \int_E Tf(s)g(s)d\mu(s) = \int\int_{E\times E} K(s,t)f(t)$ $g(s)d\mu(s)d\mu(t) = \langle f, Tg\rangle_2 \Rightarrow$ T self-adjoint. □

Lemma

$||K - K_n||_{sup} \to 0 \Rightarrow ||T - T_n|| \to 0.$

Proof.

$||(T_n - T)f||_2^2 = \int_E(\int_E(K(s,t) - K_n(s,t))f(t)d\mu(t))^2 d\mu(s)$ $\leq ||K - K_n||_{sup}^2 ||f||_2^2 \; \mu(E).$ □

Theorem

K is non-negative definite $\Leftrightarrow$ T is non-negative.

Proof.

"$\Rightarrow$": $\forall n > 0$, $\exists$ $\{v_n\}$ s.t. $E \times E \subset \cup_{m,k} B((v_m, v_k); 1/n)$, define $K_n(s,t) = K(v_m, v_k)$, if $(s,t) \in B((v_m, v_k); 1/n)$. Then $||K_n - K||_{sup} \to 0$, and $\langle T_n f, f \rangle_2 = \int \int_{E \times E} K_n(s,t) f(s) f(t) d\mu(s) d\mu(t) = \sum_{m,k} K(v_m, v_k) a_m a_k \geq 0 \Rightarrow \langle Tf, f \rangle_2 \geq 0$.

"$\Leftarrow$": If $\sum_{n,m} K(z_n, z_m) a_n a_m < 0$, then $\exists$ a disjoint $\{E_n\}$ s.t. $v_n \in E_n$ and $\max_{v_n \in E_n, v_m \in E_m} \sum_{n,m} K(v_n, v_m) a_n a_m < 0$, $\sum_{n,m} a_n a_m (\mu(E_n)\mu(E_m))^{-1} \int_{E_n} \int_{E_m} K(s,t) d\mu(s) d\mu(t) < 0 \Rightarrow$ $\langle Tf, f \rangle < 0$, $f = \sum_n a_n I_{E_n} / \mu(E_n)$. □

Theorem

$$f \in L^2(E),\ T(f) = \int_E K(s,t)f(t)d\mu(t) \Rightarrow T \in K(L^2[E]).$$

Proof.

E is compact, for a kernels K are continuous on $E \times E$, $\exists$ a degenerate kernel sequence $\{K_n\}$ s.t. $||K - K_n||_{sup} \to 0$.

Let $K_n = \sum_{n_k} g_{n_k}(s)h_{n_k}(t)$, finite n_k and $g_{n_k}, h_{n_k} \in C(E)$. $T_nf(s) = \int_E \sum_{n_k} g_{n_k}(s)h_{n_k}(t)f(t)d\mu(t) = \sum_{n_k} a_{a_k}g_{n_k}(s)$, then $Rank(T_n) < \infty \Rightarrow T_n \in K(L^2(E)) \Rightarrow T \in K(L^2(E))$. $\square$

T has a svd decomposition: $T = \sum_n \lambda_n e_n \otimes e_n$. $Te_n = \lambda_n e_n$: $e_n(s) = \int_E K(s,t)e_n(t)d\mu(t)/\lambda_n$.

Theorem

$\sum_n \lambda_n e_n(s) e_n(t) \rightrightarrows K(s,t)$

Proof.

Let $s = t$, $\sum_n \lambda_n e_n^2(t) = \sum_n (\int_E K(t,t) e_n(t) d\mu(t)) e_n(t) = \sum_n \langle K, e_n \rangle_2 e_n(t) \leq K(t,t)$.

$|\sum_n \lambda_n e_n(s) e_n(t)| \leq |\sum_n \lambda_n e_n^2(s)|^{1/2} |\sum_n \lambda_n e_n^2(t)|^{1/2} \leq (K(s,s)K(t,t))^{1/2}$.

Let $H(s,t) = \sum_n \lambda_n e_n(s) e_n(t)$. Then $T_H = \sum_n \lambda_n e_n \otimes e_n$ since $T_H f = \int \sum_n \lambda_n e_n(s) e_n(t) f(t) d\mu(t) = \sum_n \lambda_n e_n \langle e_n, f \rangle_2 \Rightarrow T = T_H \Rightarrow H = K$.

Let $g_k(s) = \sum_{n \geq k} \lambda_n e_n^2(s)$, $\forall \varepsilon > 0$, $E_k = \{g_k < \varepsilon\}$. Then $E_k \subset E_{k+1} \Rightarrow \cup_k E_k$ is open cover of $E \Rightarrow g_k \rightrightarrows 0$(Dini). □

Theorem

K non-negative symmetric kernel, $tr(T) = \int_E K(s,s)d\mu(s)$, $||T||^2_{HS} = \int\int_{E\times E} K^2(s,t)d\mu(s)d\mu(t)$.

Proof.

$tr(T) = \sum_n \lambda_n = \sum_n \lambda_n \int_E e_n^2(s)d\mu(s) = \int_E(\sum_n \lambda_n e_n^2(s))$ $d\mu(s) = \int_E K(s,s)d\mu(s)$.

$||T||_{HS} = \sum_n \lambda_n^2 = \sum_n \lambda_n^2 \int_E e_n^2(s)d\mu(s) \int_E e_n^2(t)d\mu(t) =$ $\int\int \sum_n \lambda_n^2 e_n^2(t)e_n^2(s)d\mu(s)d\mu(t) = \int\int_{E\times E}(\sum_n \lambda_n e_n(s)e_n(t))^2$ $d\mu(s)d\mu(t) = \int\int(K(s,t))^2 d\mu(s)d\mu(t)$. $\square$

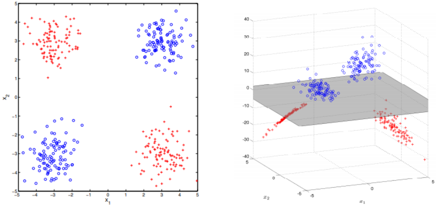
Property

$\int\int_{E\times E}(K(s,t)-H(s,t))^2 d\mu(t)d\mu(s) \geq \sum_{k>n}\lambda_n^2$, $\forall H$ s.t. $Rank(H)=n$.

Proof.

$\int\int_{E\times E}(K(s,t)-H(s,t))^2 d\mu(t)d\mu(s) = ||T_K - T_H||_{HS}$, when $T_H = \sum_{k\leq n}\lambda_k e_k \otimes e_k \Leftrightarrow H(s,t) = \sum_{k\leq n}\lambda_k e_k(s)e_k(t)$. □

Sometime we want to map objects from low dimensions to higher dimensions. To achieve this, we should find the map ϕ.



We just need to specified $K(s,t) := \langle \phi(s), \phi(t) \rangle$. Existence of ϕ is ensured by Moore-Aronszajn theorem.

Accurately, $\phi(s) = K(\cdot, s)$ and the high dimension space is "$\overline{span\{K(\cdot, s)\}}$".

Definition

H is a Hilbert space of functions: $E \to R$, $K : E \times E \to R$ is said to be a reproducing kernel for H if

(i) $K(\cdot, t) \in H$, $\forall t \in E$

(ii) $\forall f \in H$, and $t \in E$, $f(t) = \langle f, K(\cdot, t)\rangle$

$$K(s,t) = K(\cdot, t)|_{\cdot = s} = \langle K(\cdot, t), K(\cdot, s)\rangle.$$

Theorem

Evaluation functional: $J_t : H \to R$, $J_t(f) = f(t)$. Then H is an RKHS if $\{J_t\}_{t\in E} \subset H^*$.

Proof.

$f(t) = J_t(f) = \langle f, g_t\rangle$, let $K(s,t) = g_t(s)$. □

Example

$(V, \langle\cdot,\cdot\rangle)$ *finite dimension inner product space of functions,* $\{e_n\}$ *COB.* $|J_x(f)| = |f(x)| = |\sum_n a_n e_n(x)| \leq (\sum_n e_n^2(x))^{1/2}||f||$, *then* V *is RKHS.*

Let $K(s,t) = \sum_n e_n(s)e_n(t)$, *then* $K(\cdot,t) \in V$, $\langle f, K(\cdot,t)\rangle$ $= \langle \sum_m b_m e_m, \sum_n e_n(t)e_n(\cdot)\rangle = \sum_n \langle b_n e_n(t)e_n, e_n\rangle = \sum_n b_n e_n(t)$ $= f(t)$.

Example

$f \in C([0,1])$, $\sup_{||f||_{sup}=1} |J_t(f)| = |f(x)| \leq 1$ *but* $C([0,1])$ *is not Hilbert space.* $f \in L^2$, $\sup_{||f||_2=1} |J_t(f)| = |f(x)|$, *but* f *may not be bounded.*

Property

H Hilbert space contained function on E with rk K
(i) K is a symmetric and non-negative definite kernel.
(ii) Reproducing kernel of H is unique.

Proof.

(i) $\sum_i \sum_j a_i a_j K(t_i, t_j) = \sum_i \sum_j a_i a_j \langle K(\cdot, t_i), K(\cdot, t_j)\rangle = \langle \sum_i a_i K(\cdot, t_i), \sum_j a_j K(\cdot, t_j)\rangle \geq 0$

(ii) Let K_1, K_2 be two reproducing kernel, then $\forall f \in H$, $f(t) = \langle f, K_1(\cdot, t)\rangle = \langle f, K_2(\cdot, t)\rangle \Rightarrow \langle f, K_1(\cdot, t) - K_2(\cdot, t)\rangle = 0$, $\forall t \in E$, $f \in H \Rightarrow K_1(\cdot, t) - K_2(\cdot, t) = 0$, $\forall t \Rightarrow K_1 = K_2$. $\square$

Definition

K is a symmetric and non-negative definite kernel of set E,
$H_0 = span\{k(\cdot,t), t \in E\} = \{\sum_n a_n K(\cdot,t_n); a_n \in R, t_n \in E\}$
$\langle \sum_n a_n K(\cdot,t_n), \sum_m b_m K(\cdot,s_m)\rangle = \sum_n \sum_m a_n b_m K(t_n,s_m)$.

Theorem

$(H_0, \langle\cdot,\cdot\rangle)$ is an inner product space.

Proof.

H_0 is a vector space over field R. And $\langle\cdot,\cdot\rangle$ is a bilinear non-negative symmetric function. To establish $\langle\cdot,\cdot\rangle$ is an inner product, let $f \in H_0$, if $\langle f, f\rangle = 0$, $|f(t)| = |\langle f, K(\cdot,t)\rangle| \leq ||f||$ $K^{1/2}(t,t) \Rightarrow f = 0$. □

Theorem

$||f_n - f|| \to 0 \Rightarrow f_n(t) \to f(t), \forall t \in E$.

If $\{f_n\}$ is Cauchy in H_0, then $\exists$ pointwise limit f. And if $f \in H_0$, then $||f_n - f|| \to 0$.

Proof.

$|f_n(t) - f(t)| = |\langle f_n - f, K(\cdot, t)\rangle| = ||f_n - f||\ K^{1/2}(t, t)$

Let $\{f_n\}$ Cauchy, $\{f_n(x)\}$ Cauchy since $|f_n(x) - f_m(x)|$ $= |\langle f_n - f_m, K(\cdot, x)| \le ||f_n - f_m||K^{1/2}(x, x)$, let $f_n(x) \to y_x$.

Define $f : E \to R$, $f(x) = y_x$. $||f_n - f|| \to 0$ since that

$||f_m - f_n||^2 = ||f_m - f||^2 + ||f - f_n||^2 - 2\langle f_m - f, f_n - f\rangle$

$\langle f_m - f, f_n - f\rangle = \langle f_m - f, \sum_j a_j K(\cdot, t_j)\rangle = \sum_j a_j(f_m(t_j) - f(t_j))$

$\Rightarrow \limsup_m ||f_m - f_n||^2 = \limsup_m ||f_m - f||^2 + ||f - f_n||^2$

$\Rightarrow ||f - f_n||^2 \le \limsup_m ||f_m - f_n||^2$. $\square$

Definition

$H(K) = H_0 \cup \{$*the pointwise limits of all the Cauchy in* $H_0\}$
If $f_n \to f$, $g_n \to g$, *define* $\langle f, g \rangle = \lim_n \langle f_n, g_n \rangle$

We ignore some details about the validation of the definition since this is similar to the completion of metric spaces.

Theorem

$(H(K), \langle \cdot, \cdot \rangle)$ *is a Hilbert space.*

Property

$(H(K), \langle \cdot, \cdot \rangle)$ is the unique Hilbert space of functions on E with a symmetric non-negative kernel K as its reproducing kernel.

Proof.

Supposed that $(G, \langle \cdot, \cdot \rangle_G)$ is another Hilbert space of functions on E with a symmetric non-negative kernel K as its reproducing kernel. $K(\cdot, t) \in G \Rightarrow H(K) \subset G$, then $H(K)$ is closed sub-vector space of G, then $G = H(K) \oplus H(K)^{\perp}$. $\forall$ $f \in H(K)^{\perp}$, $f(t) = \langle f, K(\cdot, t) \rangle_G = 0 \Rightarrow f = 0$ □

Property

If K is continuous on $E \times E$, then $H(K) \subset C(E)$.

Moreover, If E is separable metric space, then $H(K)$ is separable.

Proof.

$\forall x \in E$, $f \in H(K)$, $|f(x) - f(y)| = |\langle f, K(\cdot, x) - K(\cdot, y)\rangle|$ $\leq ||f||\ ||K(\cdot,x) - K(\cdot,y)|| = ||f||\ (K(x,x) - 2K(x,y) + K(y,y))^{1/2}$

$H_1 = \{\sum_n a_n K(\cdot, t_n); a_n \in Q, t_n \in E\}$, H_1 is dense in $H_0 \Rightarrow$ H_1 is dense in $H(K)$. $\square$

$K(s,t) = \langle K(\cdot, s), K(\cdot, t)\rangle = \langle \sum_n e_n(t) e_n, \sum_m e_m(s) e_m\rangle = \sum_n e_n(t) e_n(s)$. Note that $\sum_n e_n(t) e_n(s) \rightrightarrows K(s,t)$.

Theorem

$t \in E$, $g(t,\cdot) \in L^2(S,\mathcal{F},\mu)$, $K(t,t^{'}) = \int g(t,s)g(t^{'},s)d\mu(s)$. *Then* $H(K) = \{\int F(s)g(t,s)d\mu(s); F \in G\}$, $G = \overline{span\{g(t,\cdot)\}}$ *and* $H(K) \approxeq G$.

Proof.

$\langle g(t,\cdot), g(t^{'},\cdot)\rangle_2 = K(t,t^{'}) = \langle K(t,\cdot), K(t^{'},\cdot)\rangle$, and $G \subset L^2$ $\Rightarrow H(K) \approxeq G$. □

Let $\underline{X_t \in L^2(E)}$ and $\underline{EX_t} = 0$, then $K(t,s) = EX_tX_s$. We have $H(K) \approxeq \overline{span\{X_t\}}$, $\overline{span\{X_t\}}$ is called the closed span of $\{X_t\}$, we mark that $L^2(X)$.

Example

Let $K(s,t) = min(s,t) = \int I_{[0,s]}(x) I_{[0,t]}(x)dx$, which is the covariance of Brown motion. $H(K) = \{\int_0^x F(t)dt,\ F \in L^2([0,1])\}$.

Theorem

$f \in H(K) \Leftrightarrow \exists\ Y \in L^2(X)$ s.t. $f(t) = E(YX_t)$.

Proof.

"$\Rightarrow$": $f(t) = \langle f, K(\cdot,t)\rangle = EYX_t$.

"$\Leftarrow$": Let $Y_n = \sum_{n_k} a_{n_k} X_{t_{n_k}} \to Y$. Let $\sum_{n_k} a_{n_k} K(t_{n_k},\cdot) \to f(\cdot)$, $E(Y_n X_t) = \sum_{n_k} a_{n_k} K(t_{n_k}, t) \to f(t)$.

$|EY_nX_t - EYX_t| \le ||Y_n - Y||_2 ||X_t||_2 \Rightarrow EY_nX_t \to EYX_t \Rightarrow f(t) = EYX_t$. $\square$

Let K_1, K_2 be two symmetric non-negative kernels, and $K = K_1 + K_2$. We want to prove $H(K) = H(K_1) + H(K_2)$.

Let $F = H(K_1) \times H(K_2)$ equipped with inner product $\langle (f_1, f_2), (g_1, g_2) \rangle_F = \sum_i \langle f_i, g_i \rangle_{K_i}$. Noticed that F is Hilbert space.

Let $F_1 = \{(f, -f); f \in \cap_i H(K_i)\}$ equipped with $\langle \cdot, \cdot \rangle_F$. We can prove that F_1 is a closed sub-space of F.

Then $F = F_1 \oplus F_1^{\perp} \Rightarrow \forall\ h = f_1 + f_2$, $\exists (g_1, g_2) \in F_1^{\perp}$ and $f \in \cap_i H(K_i)$ s.t. $h = f_1 + f_2 = f + g_1 + (-f) + g_2 = g_1 + g_2$.

Define $\Phi : H(K_1) + H(K_2) \to F_1^{\perp}$, then Φ is a invertible linear map. $u, v \in H(K_1) + H(K_2)$, define $\langle u, v \rangle_K = \langle \Phi(u), \Phi(v) \rangle_F$. Then $H(K_1) + H(K_2)$ is Hilbert space.

Lemma

$$H(K) = H(K_1) + H(K_2).$$

Proof.

"$\subset$": $\sum_n a_n K(\cdot, t_n) = \sum_n a_n K_1(\cdot, t_n) + \sum_n a_n K_2(\cdot, t_n)$.

"$\supset$": $\forall\ h \in \sum_i H(K_i)$, $h(t) = f_1(t) + f_2(t) = g_1(t) + g_2(t) = \langle g_1, K_1(\cdot, t)\rangle_{K_1} + \langle g_2, K_2(\cdot, t)\rangle_{K_2} = \langle (g_1, g_2), (K_1(\cdot, t), K_2(\cdot, t))\rangle_F = \langle (g_1, g_2), P_{F_1^\perp}(K_1(\cdot, t), K_2(\cdot, t))\rangle_F = \langle h, K(\cdot, t)\rangle_K$. □

$$\forall f \in H(K),\ ||f||_K^2 = ||(g_1, g_2)||_F^2 = ||g_1||_{K_1}^2 + ||g_2||_{K_2}^2.$$

Theorem

$$\forall f \in H(K),\ ||f||_K^2 = \min_{f_i \in H(K_i), f = f_1 + f_2}(||f_1||_{K_1}^2 + ||f_2||_{K_2}^2).$$

Proof.

$||f_1||_{K_1}^2 + ||f_2||_{K_2}^2 = ||(f_1, f_2)||_F^2 = ||(f_1 - g_1, f_2 - g_2) + (g_1, g_2)||_F^2 = ||g_1||_{K_1}^2 + ||g_2||_{K_2}^2 + ||(f_1 - g_1, f_2 - g_2)||_F^2 \geq ||f||_K^2.$ □

Definition

If K_1, K_2 are two symmetric kernels, and $K_2 - K_1$ is non-negative, then $K_1 \ll K_2$.

Theorem

Let K_1, K_2 be two symmetric non-negative kernels, and $\exists B > 0$, $K_1 \ll BK_2$, then $H(K_1) \subset H(K_2)$.

Proof.

$$BK_2 = K_1 + K_3, \forall f \in H(K_1), B^2||f||_{K_2}^2 \leq ||f||_{K_1}^2. \qquad \square$$

Definition

H_1 and H_2 are separable Hilbert spaces of functions on E, $\{e_{in}\}$ be COB of H_i, let $H = H_1 \otimes H_2 = \{\sum_{j,g} a_{jg} e_{1j}(s) e_{2g}(t)\}$, $\langle a, b \rangle = \sum_{j,g} a_{jg} b_{jg}$. H is a Hilbert space.

Noticed that $H_1 \times H_2 = \{(f, g); f \in H_1, g \in H_2\}$. $H_1 \otimes H_2$ just $\{fg : f \in H_1, g \in H_2\}$ with $\langle f_1 g_1, f_2 g_2 \rangle = \langle f_1, g_1 \rangle_1 \langle f_2, g_2 \rangle_2$.

Theorem

$H(K_1) \otimes H(K_2)$ is RKHS with rk $K_1 K_2$.

Proof.

$$|f(s,t)| = |\sum_{j,g} a_{jg} e_{1j}(s) e_{2g}(t)| = |\sum_j e_{1j}(s) \ \sum_g a_{jg} e_{2g}(t)|$$
$$\leq |\sum_j e_{1j}(s) (\sum_g a_{jg}^2)^{1/2} \ |(\sum_g e_{2g}^2(t))^{1/2}$$
$$\leq (\sum_j e_{1j}^2(s))^{1/2} \ ||f||^2 \ (K_2(t,t))^{1/2}$$
$$= (K_1(s,s))^{1/2} \ ||f||^2 \ (K_2(t,t))^{1/2} \Rightarrow H \text{ is an RKHS.}$$

$$f(s,t) = \sum_{j,g} a_{jg} e_{1j}(s) e_{2g}(t) =$$
$$\sum_{j,g} a_{jg} \langle e_{1j}, K_1(\cdot, s)\rangle_1 \langle e_{2g}, K_2(*, t)\rangle_2 =$$
$$\sum_{j,g} a_{jg} \langle e_{1j} e_{2g}, K_1(\cdot, s) K_2(*, t)\rangle = \langle f, K_1(\cdot, s) K_2(*, t)\rangle.$$

$\square$

Definition

For $T \in B(H(K))$, $(Tf)(s) = \langle Tf, K(\cdot, s)\rangle = \langle f, T^*K(\cdot, s)\rangle$, define a kernel of T: $R(s,t) = T^*K(\cdot, s)|_{\cdot = t}$.

The existence of kernel for T since that $J_s = Tf(s) = \langle f, e_s\rangle$.
R is symmetric $\Leftrightarrow$ T is self-adjoint.
$aT + bG \Rightarrow aR_T + bR_G$.
$T = HG$, $R_T(s,t) = G^*H^*K(\cdot, s)|_{\cdot = t} = G^*R_H(\cdot, s)|_{\cdot = t}$
$= \langle G^*R_H(\cdot, s), K(*, t)\rangle = \langle R_H(\cdot, s), R_{G^*}(*, t)\rangle$.

Definition

$T \in B(H(K))$ is non-negative $\Leftrightarrow$ R_T is non-negative.

Proof.

$"\Rightarrow"$: $\sum_{n,m} a_n a_m R_T(t_n, t_m) = \sum_{n,m} a_n a_m T^* K(\cdot, t_m)|_{\cdot = t_n}$
$= \sum_{n,m} a_n a_m \langle T^* K(\cdot, t_m), K(*, t_n)\rangle = \langle f, Tf\rangle \geq 0.$
$"\Leftarrow"$: $\langle T^* f, f\rangle = \sum_{n,m} a_n a_m R_T(t_n, t_m)$. □

If T non-negative, $0 \leq \langle Tf, f\rangle \leq \langle ||T||f, f\rangle$, then $||T||I - T$ is non-negative, $(||T||I - T^*)K(\cdot, t)|_{\cdot = s}$ non-negative.

Definition

For $T \in B(H(K_1), H(K_2))$, $R_T(s, t) = T^* K_2(\cdot, s)|_{\cdot = t}$.

Theorem

$H(K_i)$ separable Hilbert space, $T \in B_{HS}(H(K_1), H(K_2)) \Leftrightarrow \sum_{i,j} a_{ij}^2 < \infty$, $R(s,t) = \sum_{i,j} a_{ij} e_{1i}(s) e_{2j}(t)$.

Proof.

$"\Rightarrow"$: $R(s,t) = T^* K_2(\cdot,t)|_{\cdot=s} = \sum_i \langle T^* K_2(\cdot,t), e_{1i}\rangle_1 e_{1i}(s) = \sum_i \langle K_2(\cdot,t), \sum_j \langle T e_{1i}, e_{2j}\rangle_2 e_{2j}\rangle_2 e_{1i}(s) = \sum_{i,j} a_{ij} e_{2j}(t) e_{1i}(s)$

$"\Leftarrow"$: $||Te_{1n}||^2 = ||\langle e_{1n}, R(\cdot,s)\rangle||^2 = ||\sum_j a_{nj} e_{2j}||^2 \Rightarrow ||T||_{HS}^2 = \sum_{i,j} a_{ij}^2 < \infty$. $\square$

$$tr(T) < \infty \Leftrightarrow \sum_{ij} |a_{ij}| < \infty.$$

$\forall f \in C_q[0,1]$, a taylor expansion of $f(t)$:

$$\sum_{k\le q-1} f^{(k)}(0)\frac{t^k}{k!} + \int_0^t f^{(q)}(u)\frac{(t-u)^{q-1}}{(q-1)!}du$$

Let $\Phi_k(t) = \frac{t^k}{k!}$, $G_q(u) = \frac{u_+^{q-1}}{(q-1)!}$, we rewrite the taylor expansion of $f(t)$:

$$\sum_{k\le q-1} \frac{f^{(k)}(0)}{\Phi_k(t)} + \int_0^1 f^{(q)}(u)G_q(t-u)du$$

Definition

$f \in W_q[0,1] : f(t) = \sum_{k=0}^{q-1} a_k\Phi_k(t) + (G_q \circ g)(t)$, $g \in L^2[0,1]$.

Property

$f \in W_q[0,1]$, $f^{(j)}(t) = \sum_{k=j}^{q-1} a_k \Phi_{k-j}(t) + (D^{(j)} G_q \circ g)(t)$, $j < q$ and $f^{(q)} = g$.

$\forall\ f = G_q \circ g$, $f^{(i)}(0) = 0$, $f^{(q)} = g$.

Definition

$H_0 = span\{\Phi_i\}_{i=0}^{q-1} = P_q[t]$, $\langle \alpha, \beta \rangle_{H_0} = \sum_i \alpha^{(i)}(0)\beta^{(i)}(0)$. Then $(H_0, \langle \cdot, \cdot \rangle_{H_0})$ is a Hilbert space and $\{\Phi_i\}$ is COB of H_0.

$H_1 = \{G_q \circ g; g \in L^2[0,1]\}$, $\langle f, h \rangle_{H_1} = \int_0^1 f^{(q)}(u)h^{(q)}(u)du$. Then $(H_1, \langle \cdot, \cdot \rangle_{H_1})$ is a Hilbert space since $H_1 \cong L^2[0,1]$.

Theorem

$K_0(s,t) = \sum_i \Phi_i(s)\Phi_i(t)$ *is rk of* H_0.
$K_1(s,t) = \int_0^1 G_q(s-u)G_q(t-u)du$ *is rk of* H_1

Proof.

$H_1 = \{\int_0^1 g(u)G_q(t-u)du; g \in L^2(E)\}$ and $L^2(E) = \overline{span\{G_q(\cdot - u)\}}$. □

Theorem

$W_q[0,1]$ *is a RKHS with rk* $K_0 + K_1$. *Since* $H_0 \cap H_1 = \{0\}$, $\langle f, g\rangle_{W_q} = \langle f_1 + f_2, g_1 + g_2\rangle_{W_q} = \langle f_1, g_1\rangle_{H_0} + \langle f_2, g_2\rangle_{H_1} \Rightarrow W_q[0,1] = H_0 \oplus H_1$.

Definition

Define another inner product $\langle \cdot, \cdot \rangle_{2,q}$*:* $\langle f, g \rangle_{2,q} = \langle f, g \rangle_2 + \langle f^{(q)}, g^{(q)} \rangle_2$

$|| \cdot ||_{2,q}$ *and* $|| \cdot ||_{W_q}$ *are equivalent norms.*

Proof.

$||f||_{W_q}^2 = ||\sum_n b_n \Phi_n||_{H_0}^2 + ||G_q \circ f^{(q)}||_{H_1}^2 = \sum_n b_n^2 + ||f^{(q)}||_2^2$
and $||f||_{2,q}^2 = ||f||_2^2 + ||f^{(q)}||_2^2$

$||f||_2^2 = ||\sum_n b_n \Phi_n + G_q \circ f^{(q)}||_2^2 \le C_1 (\sum_n |b_n| + ||f^{(q)}||_2)^2$
$\le C_2 ||f||_{W_q}^2 \Rightarrow ||f||_{2,q}^2 = ||f||_2^2 + ||f^{(q)}||_2^2 \le C_2 ||f||_{W_q}^2 + ||f^{(q)}||_2^2$
$\le (C_2 + 1) ||f||_{W_q}^2$. □

Under $\langle \cdot, \cdot \rangle_{2,q}$, we apply Gram-Schmidt to $\{\Phi_n(t)\}_{n=0}^{q-1}$ and get the Legendre polynomials $\{p_n(t)\}_{n=0}^{p-1}$.

Lemma

Exist a COB $\{e_n\}$ for $L^2([0,1])$ s.t. $\langle e_i^{(q)}, e_j^{(q)} \rangle_2 = \gamma_i \delta_{ij}$ and $\gamma_1 = ... = \gamma_q = 0$, $C_1 j^{2q} \leq \gamma_{j+q} \leq C_2 j^{2q}$.

Then $\langle e_i, e_j \rangle_{2,q} = \langle e_i, e_j \rangle_2 + \langle e_i^{(q)}, e_j^{(q)} \rangle_2 = (1 + \gamma_i)\delta_{ij}$.

Theorem

$\{e_i(1+\gamma_i)^{-1/2}\}$ is a COB of $W_q([0,1])$ and $\{e_n\}_{n=0}^{p-1}$ is Legendre polynomials.