Stat 860, Lecture 4
(Do reading from hw1, and hw2 by Lecture 6)

To define a particular Hilbert space $\mathcal{L}_2\Big[\{\phi_\nu\}, \Lambda\Big]$:
$\{\phi_\nu\}$an orthonormal sequence of elements in $\mathcal{L}_2[0,1]$:

$$(\phi_\nu, \phi_\mu) \equiv \int_0^1 \phi_\nu(s)\phi_\mu(s)ds = \begin{cases} 1 & \text{if } \nu = \mu \text{ ,} \\ 0 & \text{otherwise.} \end{cases}$$

$\Lambda = (\lambda_1, \lambda_2, \ldots) \qquad \lambda_\nu > 0$

$$\mathcal{L}_2\Big[\{\phi_\nu\}, \Lambda\Big] \;:=\quad f \;:\; \sum_\nu \frac{(f, \phi_\nu)^2}{\lambda_\nu} < \infty$$

Λ: We will only be interested in the case $\lambda_\nu \to 0$ as $\nu \to \infty$

$$\langle f, g \rangle_{\mathcal{L}[\{\phi_\nu\},\Lambda]} = \sum_\nu \frac{(f, \phi_\nu)(g, \phi_\nu)}{\lambda_\nu}$$

If $\lambda_\nu \to 0$, then $\mathcal{L}_2[\{\phi_\nu\}, \Lambda] \subset \mathcal{L}_2[0,1]$, but different geometry.

$\mathcal{H}$ a Hilbert space

L a bounded linear functional means

$$L : f \to \text{Real numbers}$$
$$L(f_1 + f_2) = L\,f_1 + L\,f_2$$

Notation: "$\exists M \ni$" means "there exists M such that".

"bounded" means

$$\exists M \ni |Lf| \leq M\|f\|,$$

equivalently

$$\sup_{\|f\| \leq 1} |Lf| < \infty$$

Riesz Representation Theorem: If L is a bounded linear functional in $\mathcal{H}$, then $\exists \eta \in \mathcal{H} \ni$

$$Lf = \langle \eta, f \rangle \qquad \text{all } f \in \mathcal{H}$$

η is called the *representer* of L.

Example: $\mathcal{L}_2[0,1]$

If $\eta \in \mathcal{L}_2[0,1]$, then

$$Lf = (\eta, f) = \int_0^1 \eta(s) f(s) ds$$

is a bounded linear functional on $\mathcal{L}_2[0,1]$.

For all (symbol: $\forall$) L, there is an η and conversely.

$Lf = (\eta, f)$ is a bounded linear functional:

$$(\eta, f_1 + f_2) = (\eta, f_1) + (\eta, f_2)$$

$$(\eta, f) \leq M\|f\| \quad \text{with } M = \|\eta\|$$

(Cauchy-Schwarz Inequality: $|(f,g)| \leq \|f\| \, \|g\|$)

Important remark:

Let $Lf \equiv L_{t_*}f = f(t_*)$ for some fixed $t_* \in [0,1]$.

L is NOT a bounded linear functional on $\mathcal{L}_2[0,1]$—you cannot find an M so that

$$|Lf| \leq M\sqrt{\int_0^1 f^2(\mu)d\mu}.$$

(In fact, for $f \in \mathcal{L}_2,\ f(t_*)$ may not be uniquely defined: elements in $\mathcal{L}_2[0,1]$ are really only equivalence classes of functions that are equal almost everywhere!)

That's why solving variational problems in $\mathcal{L}_2[0,1]$ is usually not very interesting.

Definition Let $\mathcal{T}$ be an index set (for example, $\mathcal{T} = [0, 1]$). A (real) *Reproducing Kernel Hilbert Space* (RKHS) is a Hilbert space of real-valued functions defined on $\mathcal{T}$ for which all the evaluation functionals are bounded linear functionals.

An evaluation functional: let $t_* \in \mathcal{T}$. Then

$$Lf = f(t_*)$$

is an evaluation functional at t_*. L will be bounded if

$$|Lf| \equiv |f(t_*)| \leq M_{t_*} \|f\|$$

for some constant M_{t_*} (not depending on f).

Example:

Let $\phi_1, \phi_2, \ldots$ be an orthonormal set (o.n.s.) in $\mathcal{L}_2[0,1]$ with a continuous representation; for example

$$\{1, \sqrt{2}\sin 2\pi\nu t, \sqrt{2}\cos 2\pi\nu t,\ \nu = 1, 2, \ldots\}$$

Let $\lambda_\nu,\ \nu = 1, 2 \ldots > 0$ and suppose $\{\lambda_\nu, \phi_\nu\}$ such that

$$\sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu^2(t) \leq M < \infty, \quad \text{all } t \in [0,1]$$

(since $|\sin 2\pi\nu t|,\ |\cos 2\pi\nu t| \leq 1$, it is sufficient that

$$\sum_{\nu=1}^{\infty} \lambda_\nu < \infty$$

if the $\{\phi_\nu\}$ are sines and cosines)

Theorem Then $\mathcal{L}_2[\{\phi_\nu\}, \lambda_\nu]$ is an RKHS.

Proof. We have to find an M_t such that

$$|f(t)| \le M_t \|f\| \qquad \text{for each } t \in \mathcal{T}$$

Recall

$$\|f\|^2 = \sum_{\nu=1}^{\infty} \frac{(f, \phi_\nu)^2}{\lambda_\nu}$$

$$f \in \mathcal{L}_2[\{\phi_\nu\}, \lambda_\nu]$$

Since $\lambda_\nu \to 0,\ f \sim \sum (f, \phi_\nu)\phi_\nu$

$$\begin{aligned}
|f(t)| &= \left| \sum_{\nu=1}^{\infty} (f, \phi_\nu)\phi_\nu(t) \right| \\
&\le \sum_{\nu=1}^{\infty} |(f, \phi_\nu)|\,|\phi_\nu(t)| \\
&\equiv \sum_{\nu=1}^{\infty} \frac{|(f, \phi_\nu)|}{\sqrt{\lambda_\nu}} \cdot \sqrt{\lambda_\nu}|\phi_\nu(t)| \\
&\le \sqrt{\sum_{\nu=1}^{\infty} \frac{|(f, \phi_\nu)|^2}{\lambda_\nu}} \sqrt{\sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu^2(t)}
\end{aligned}$$

$$\sqrt{\sum_{\nu=1}^{\infty} \frac{|(f,\phi_\nu)|^2}{\lambda_\nu}} \sqrt{\sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu^2(t)} \leq \|f\| \quad M$$

So

$$|f(t)| \leq M\|f\|, \quad all \;\; t$$

Remark: Recall $E_n(\Sigma)$ $\quad \Sigma = \Gamma D \Gamma'$

$$\langle x, y \rangle = x'\Sigma^{-1}y = \sum_{\nu=1}^{n} \frac{(x,\phi_\nu)(y,\phi_\nu)}{\lambda_\nu}$$

Let $\{\phi_\nu\}$ be the columns of Γ,

$$\phi_\nu = (\phi_\nu(1), \ldots, \phi_\nu(n))'$$

$$\Sigma_{ij} = \sum_{\nu=1}^{n} \lambda_\nu \phi_\nu(i)\phi_\nu(j)$$

Returning to $\mathcal{L}_2[\{\phi_\nu\}, \lambda_\nu]$. If you suspect that there is something interesting about

$$K(s,t) \equiv \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(s)\phi_\nu(t)$$

you are right!

-LATER-

Let $\mathcal{H}$ be an RKHS. Then $Lf = f(t)$ is a bounded linear functional for each $t \in \mathcal{T}$. So $\exists$ a representer, call it $\xi_t \ni$

$$\langle f, \xi_t \rangle = f(t) \qquad (t \text{ fixed})$$

Claim: For $\mathcal{L}_2[\{\phi_\nu\}, \lambda_\nu]$

Fix $t = t_*$. Then

$$\xi_{t_*}(s) = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(t_*) \phi_\nu(s)$$

To show ξ_{t_*} is the representer of

$$Lf \to f(t_*)$$

we need to show

1. $\xi_{t_*} \in \mathcal{H}$

2. $\langle \xi_{t_*}, f \rangle = f(t_*) \qquad \text{all } f \in \mathcal{H}$

Proof:

$$\xi_{t_*}(\cdot) = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(t_*)\phi_\nu(\cdot)$$

So

$$(\xi_{t_*}, \phi_\nu) = \lambda_\nu \phi_\nu(t_*)$$

"Generalized Fourier coefficient"

Need

$$\sum_{\nu=1}^{\infty} \frac{(\xi_{t_*}, \phi_\nu)^2}{\lambda_\nu} = \sum_{\nu=1}^{\infty} \frac{(\lambda_\nu \phi_\nu(t_*))^2}{\lambda_\nu} \equiv \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu^2(t_*) \leq M$$

OK, so $\xi_{t_*} \in \mathcal{H}$

Is $\langle \xi_{t_*}, f \rangle = f(t_*)$?? Let $f_\nu = (f, \phi_\nu)$. Then

$$\begin{aligned}
\langle \xi_{t_*}, f \rangle &= \sum_{\nu=1}^{\infty} \frac{(f, \phi_\nu)\,(\xi_{t_*}, \phi_\nu)}{\lambda_\nu} \\
&= \sum_{\nu=1}^{\infty} \frac{f_\nu \cdot \lambda_\nu \phi_\nu(t_*)}{\lambda_\nu} \\
&= \sum_{\nu=1}^{\infty} f_\nu \phi_\nu(t_*) \\
&= f(t_*) \quad QED
\end{aligned}$$

Claim:

$$\langle \xi_s, \xi_t \rangle = \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(s) \phi_\nu(t)$$

$$\xi_s(\cdot) = \sum_{\nu=1}^{\infty} [\lambda_\nu \phi_\nu(s)] \phi_\nu(\cdot), \xi_t(\cdot) = \sum_{\nu=1}^{\infty} [\lambda_\nu \phi_\nu(t)] \phi_\nu(\cdot)$$

$$\langle \xi_s, \xi_t \rangle = \sum_{\nu=1}^{\infty} \frac{[\lambda_\nu \phi_\nu(s)][\lambda_\nu \phi_\nu(t)]}{\lambda_\nu} \equiv \sum_{\nu=1}^{\infty} \lambda_\nu \phi_\nu(s) \phi_\nu(t)$$

$$= K(s, t)$$

ABSTRACT STATEMENT (half of the Moore-Aronszajn Theorem): Let $\mathcal{H}$ be an RKHS. Then there is associated a unique positive definite function $K(s,t)$ on $\mathcal{T} \times \mathcal{T}$ given by

$$K(s,t) = \langle \xi_s, \xi_t \rangle$$

where ξ_s is the representer of $Lf \to f(s)$. $K(\cdot,\cdot)$ is called the reproducing kernel (RK) for $\mathcal{H}$.

WHAT IS A POSITIVE DEFINITE FUNCTION?

$\mathcal{T}$ can be anything, $s, t \in \mathcal{T}$ (Anything you can define a positive definite function on)

$K(s,t) = K(t,s)$ is positive definite if for every $n = 1, 2, \ldots$ and $t_1, t_2, \ldots, t_n \in \mathcal{T}$ and $a_1, \ldots, a_n$

$$\sum_{i,j=1}^{n} a_i a_j K(t_i, t_j) \geq 0$$

(i.e. every $n \times n$ matrix obtained by discretizing $K(s,t)$ is nonnegative definite)

Important Remark:

Sums and (tensor) products of positive definite functions are positive definite.

Example, let $u \in [0, 1], v \in [0, 2], t = (u, v) \in$ the rectangle $[0, 1] \otimes [0, 2]$.

Let $K_1(u, u'), u, u' \in [0, 1]$ and $K_2(v, v'), v, v' \in [0, 2]$ be positive definite functions on $[0, 1] \otimes [0, 1]$ and $[0, 2] \otimes [0, 2]$ respectively

Then, letting $t = (u, v)$, let

$K_1(u, u')K_2(v, v') \equiv K(t, t')$, say

Then, $K(t, t')$ is a positive definite function with t and t' each in the rectangle $[0, 1] \otimes [0, 2]$
(This is the continuous version of the homework due next lecture!- Show that the tensor product of two positive definite matrices is positive definite)

Try putting ”Reproducing Kernel” into the Advanced Search box in Google-

2004: 9500 hits..

2005: 71,700 hits..

2006: 127,000 hits..

2007: 192,000 hits..

2009: 253,000 hits..

2010: 96,500 hits

2011: 1,010,000hits..

2012: 704,000 hits..

2013: 897,000 hits..

2014: 1,120,000 hits..

2015: 236,000 hits..

2016: 536,000 hits..

First page has several meaty articles. See Wikipedia entry for Reproducing Kernel Hilbert Space.

”Reproducing Kernel Hilbert Spaces in Probability and Statistics” Berlinet and Thomas-Agnan, other books (Amazon !)

Remark: Recall that in $\mathcal{L}_2[0,1]$, $\|f - P_n f\| \to 0$ does NOT imply $|f(t) - (P_n f)(t)| \to 0$ but it does if $\mathcal{H}$ is an RKHS.

Theorem Let f and $f_1, f_2, \ldots \in \mathcal{H}$ with

$$\lim_{n\to\infty} \|f - f_n\| \to 0$$

If $\mathcal{H}$ is an RKHS of real-valued functions on $\mathcal{T}$, then

$$\|f - f_n\| \to 0 \Rightarrow |f(t) - f_n(t)| \to 0$$

for each $t \in \mathcal{T}$.

"Norm convergence" (a.k.a "strong convergence") implies pointwise convergence

Proof. Let ξ_t be the representer of $Lf \to f(t)$. Then

$$\begin{aligned} |f(t) - f_n(t)| &= |\langle f, \xi_t\rangle - \langle f_n, \xi_t\rangle| \\ &= |\langle f - f_n, \xi_t\rangle| \qquad (Cauchy - Schwarz) \\ &\leq \|f - f_n\| \, \|\xi_t\| \to 0 \end{aligned}$$