# Stochastic Processes: Lecture 3

Andrew J. Holbrook

UCLA Biostatistics 270

Spring 2022

# Hilbert space

A Hilbert space $H$ is an inner product space that is also a complete metric space w.r.t. the distance function

$$d(x, y) = ||x - y|| = \sqrt{\langle x - y, x - y \rangle}$$

induced by the inner product $\langle \cdot, \cdot \rangle$. For $x, y \in H$, this inner product satisfies

1. symmetry, i.e., $\langle x, y \rangle = \langle y, x \rangle$,
2. linearity, i.e., $\langle ax_1 + bx_2, y \rangle = a\langle x_1, y \rangle + b\langle x_2, y \rangle$,
3. positive definiteness, i.e.,

$$\begin{aligned} \langle x, x \rangle &> 0\,, \quad x \neq 0 \\ \langle x, x \rangle &= 0\,, \quad x = 0\,. \end{aligned}$$

# $L^2$-space

For a measure space $X$, a function $f : X \to \mathbb{R}$ is square integrable if

$$\int_X f(x)^2 dx < \infty \, .$$

The set of square integrable functions on $X$ equipped with the inner product

$$\langle f_1, f_2 \rangle = \int_X f_1(x) f_2(x) dx$$

is a Hilbert space.

# RKHS

A Hilbert space is a reproducing kernel Hilbert space $H$, if for every function $f \in H$ defined on $X$, there exists a function $K : X \times X \to \mathbb{R}$, such that

1. For all $y$, as a function of $x$, $K(x, y) \in H$
2. Reproducing property: for all $x \in X$,

$$f(x) = \langle K(\cdot, x), f(\cdot) \rangle .$$

Function $K(x, y)$ is the reproducing kernel of the space $H$.

# Symmetry

Proposition 1

*A reproducing kernel $K(\cdot, \cdot)$ is symmetric.*

Proof.

$$K(x, y) = \langle K(\cdot, x), K(\cdot, y)\rangle = \langle K(\cdot, y), K(\cdot, x)\rangle = K(y, x)\,.$$

$\square$

# Uniqueness of $K$

Proposition 2

*If a reproducing kernel $K$ exists, it is unique.*

Proof.

Assume there exists another reproducing kernel $K'$. Then

$$\begin{aligned} 0 &\leq ||K(\cdot, x) - K'(\cdot, x)||^2 \\ &= \langle K(\cdot, x) - K'(\cdot, x), K(\cdot, x) - K'(\cdot, x)\rangle \\ &= \langle K(\cdot, x) - K'(\cdot, x), K(\cdot, x)\rangle - \langle K(\cdot, x) - K'(\cdot, x), K'(\cdot, x)\rangle \\ &= \langle K(x, x) - K(x, x)\rangle - \langle K'(x, x) - K'(x, x)\rangle = 0\,. \end{aligned}$$

□

# Positive definiteness

Proposition 3

*A reproducing kernel $K$ is positive definite.*

Proof.

For any $y_1, \ldots, y_n \in X$ and $a_1, \ldots, a_n$,

$$\sum_i \sum_j K(y_i, y_j) a_i a_j = \sum_i \sum_j \langle K(\cdot, y_i), K(\cdot, y_j) \rangle a_i a_j$$

$$= \langle \sum_i a_i K(\cdot, y_i), \sum_j a_j K(\cdot, y_j) \rangle$$

$$= || \sum_i a_i K(\cdot, y_i) ||^2 \geq 0 \, .$$

□

# One-to-one relationship

## Theorem 1 (Moore-Aronszajn theorem)

*To every RKHS $H$ there exists a unique symmetric, positive definite function $K(\cdot,\cdot)$. For every symmetric, positive definite function $K(\cdot,\cdot)$ there exists a unique RKHS $H$.*

For an RKHS $H$, we have shown uniqueness, symmetry and positive definiteness of $K(\cdot,\cdot)$. For a symmetric, positive definite $K(\cdot,\cdot)$, let $H_0 = \text{span}\{K_x := K(x,\cdot), x \in X\}$. Define the inner product on $H_0$

$$\langle \sum_{i=1}^{n} a_i K_{x_i}, \sum_{j=1}^{m} b_j K_{x_j} \rangle_{H_0} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i b_j K(x_i, x_j)$$

and note it is symmetric, non-degenerate and satisfies $\langle K_x, K_y \rangle_{H_0} = K(x,y)$. Let $H$ be the completion of $H_0$, having functions

$$f(x) = \sum_{i=1}^{\infty} a_i K_{x_i}(x) \quad \text{for which} \quad \lim_{n\to\infty} \sup_{p\geq 0} || \sum_{i=n}^{n+p} a_i K_{x_i} ||_{H_0}^2 .$$

# One-to-one relationship

The reproducing property holds on this completion:

$$\langle f, K_x \rangle_H = \sum_{i=1}^{\infty} a_i \langle K_{x_i}, K_x \rangle_{H_0} = \sum_{i=1}^{\infty} K(x_i, x) = f(x)\,.$$

To prove uniqueness, let $G$ be another Hilbert space on which $K$ is the reproducing kernel. Then

$$\langle K_x, K_y \rangle_G = K(x, y) = \langle K_x, K_y \rangle_H\,.$$

By linearity, $\langle \cdot, \cdot \rangle_G = \langle \cdot, \cdot \rangle_H$ on $H_0$, so $H_0 \subset G$. But $G$ is complete, so it contains $H$, the completion of $H_0$. Finally, we need to show $G \subset H$. Let $f \in G$. Because $H \subset G$ and closed, $f = f_H + f_{H^\perp}$. Then

$$f(x) = \langle K_x, f \rangle_G = \langle K_x, f_{H^\perp} \rangle_G + \langle K_x, f_H \rangle_G = \langle K_x, f_H \rangle_G = \langle K_x, f_H \rangle_H = f_H(x)\,,$$

where we use the fact that $K_x \in H$.

# Roughness of RKHS

Proposition 4

*Norm convergence implies pointwise convergence in an RKHS $H$.*

Proof.

For any sequence $f_n \in H$,

$$|f_n(x) - f(x)| = |\langle K(\cdot, x), f_n(\cdot) - f(\cdot)\rangle| \leq ||K(\cdot, x)|| \, ||f_n(\cdot) - f(\cdot)||$$

□

# Roughness of RKHS

$L^2$ is rougher than an RKHS: norm convergence does not imply pointwise convergence. Consider $L^2([0,1])$ and the function

$$g_n(x) = x^n$$

$g_n$ converges to $g(x) = 0$ in norm:

$$||g_n||^2 = \int_0^1 x^{2n} dx = \frac{1}{2n+1} \to 0\,.$$

But $g_n(1) = 1 \neq 0$.

# Mercer's theorem

Let $K(\cdot,\cdot) : X \times X \to \mathbb{R}$ be a symmetric function and define the integral operator $T_K : L^2(X) \to L^2(X)$

$$T_K f(\cdot) = \int K(\cdot, x) f(x) dx \,.$$

$T_K$ is positive definite if for all $f \in L^2$, $\langle f, T_K f \rangle > 0$.

Theorem 2
*If $K$ is continuous and $T_K$ is positive definite, then $T_K$ has eigenfunctions $\phi_i \in L^2$ ($||\phi_i|| = 1$) with eigenvalues $\lambda_i > 0$ and for all $x, y \in X$,*

$$K(x, y) = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \phi_i(y)$$

# Fourier decompositions and RKHS

Theorem 3
*Let $K(\cdot,\cdot)$ be an $L^2$ kernel. For all $f \in L^2$, define the Fourier coefficients as*

$$f_i = \int f(x)\phi_i(x)dx\,.$$

*For all $f, g \in L^2$ the inner product on the RKHS of $K$ is*

$$\langle f, g\rangle = \sum_i \frac{f_i g_i}{\lambda_i}\,,$$

*and $f \in$ RKHS if*

$$||f||^2 = \sum_i \frac{f_i^2}{\lambda_i} < \infty\,.$$

# Fourier decompositions and RKHS

Proof.
The Fourier expansions for $f(\cdot)$ and $K(x,\cdot)$ are

$$f(\cdot) = \sum_i f_i \phi_i(\cdot) \quad \text{and} \quad K(x,\cdot) = \sum_i \lambda_i \phi_i(x)\phi_i(\cdot)\,.$$

Then the above inner product satisfies

$$\langle K(\cdot,x), f(\cdot)\rangle = \sum_i \frac{f_i \lambda_i \phi_i(x)}{\lambda_i} = \sum_i f_i \phi_i(x) = f(x)$$

so $K(\cdot,x)$ is a reproducing kernel and has corresponding norm $||f||^2 = \sum_i f_i^2/\lambda_i$. □

# Karhunen-Loeve

Let $\{X(t), t \in \mathcal{T}\}$ be a zero mean, second order ($E(X_t^2) < \infty$) stochastic process. Its covariance function $E(X(s)X(t)) = K(s,t)$ is continuous.

## Theorem 4

*Assume that $\lambda_i$ and $\phi_i$ satisfy the following equation:*

$$\int_{\mathcal{T}} K(s,t)\phi_i(t)dt = \lambda_i \phi_i(s)\,,$$

*where $\{\phi_i,\, i \in \mathbb{N}\}$ are orthogonal eigenfunctions in $L^2$ and $\{\lambda_i,\, i \in \mathbb{N}\}$ are eigenvalues. Furthermore, specify*

$$\xi_i = \frac{1}{\sqrt{\lambda_i}} \int_{\mathcal{T}} X(t)\phi_i(t)dt\,.$$

*Then,*

$$X(t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i\phi_i(t)$$

*as the following holds uniformly:*

$$\lim_{p\to\infty} E\left(X(t) - \sum_{i=1}^{p} \sqrt{\lambda_i}\xi_i\phi_i(t)\right)^2 = 0$$

Theorem 4 (cont.)

*Conversely, if* $X(t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i\phi_i(t)$ *for* $\{\xi_i\} \overset{iid}{\sim} (0, 1)$*, then*

$$\int_{\mathcal{T}} K(s, t)\phi_i(t)dt = \lambda_i\phi_i(s)\,.$$

Proof.

We'll use the whiteboard for this one. □

# KL and RKHS

Let $\{X(t), t \in \mathcal{T}\}$ be a zero mean, second order $(E(X_t^2) < \infty)$ Gaussian process with covariance function $E(X(s)X(t)) = K(s,t)$. When equipped with the covariance as inner product, the space

$$H_x = \mathrm{span}\{X(t), t \in \mathcal{T}\}$$

is isometrically isomorphic to the RKHS of $K$. If $X_m, X_n \in H_X$, i.e.,

$$X_m = \sum_i a_i X(t_i)\,, \quad X_n = \sum_i b_i X(t_i)\,,$$

then $\langle X_m, X_n \rangle = E(X_m X_n)$ is a valid inner product. The spaces are isometrically isomorphic because

$$\langle X(s), X(t) \rangle = E(X(s)X(t)) = K(t,s) = \langle K(t,\cdot), K(s,\cdot) \rangle\,.$$

# KL and RKHS

## Theorem 5

*If $\{X(t), t \in \mathcal{T}\}$ is a zero mean, second order Gaussian process with covariance function $E(X(s)X(t)) = K(s,t)$, then the sample path $X(\cdot)$ a.s. does not belong to the RKHS of $K$.*

Heuristic proof: recall that (for $f, g \in L^2$) the inner product on the RKHS of $K$ is

$$\langle f, g \rangle = \sum_i \frac{f_i g_i}{\lambda_i},$$

where $\lambda_i$ are the eigenvalues of $K$. Define the truncated KL expansion

$$X_p(t) = \sum_{i=1}^{p} \sqrt{\lambda_i}\xi_i \phi_i(t),$$

and note that

$$E\left(||X_p||^2\right) = E\left(\sum_{i=1}^{p} \frac{\sqrt{\lambda_i}\xi_i\sqrt{\lambda_i}\xi_i}{\lambda_i}\right) = \sum_{i=1}^{p} E(\xi_i^2) = p \longrightarrow \infty.$$

# Truncated KL expansion

Again, the KL expansion uses coefficients $\lambda_i$ and functions $\phi_i$ that satisfy the integral equation

$$\int_{\mathcal{T}} K(s,t)\phi_i(t)dt = \lambda_i \phi_i(s)\,.$$

Theorem 6

*Among all truncated expansions that take the form*

$$X_p(t) = \sum_{i=1}^{p} \sqrt{l_i} x_i \psi_i(t)\,, \quad \textit{for} \quad \int \psi_i(t)\psi_j(t)dt = \delta_{ij}\,,$$

*the truncated KL expansion minimizes the integrated mean squared error*

$$\int E(e_p^2(t))dt\,, \quad \textit{where} \quad e_p(t) = \sum_{i>p} \sqrt{l_i} x_i \psi_i(t)\,.$$

# Truncated KL expansion

Proof. The expected squared error is

$$E(e_p^2(t)) = E\left(\sum_{i>p}\sum_{j>p}\sqrt{l_i l_j}x_i x_j \psi_i(t)\psi_j(t)\right)$$
$$= E\left(\sum_{i>p}\sum_{j>p}\psi_i(t)\psi_j(t)\int\int X(t_1)X(s_1)\psi_i(t_1)\psi_j(s_1)ds_1 dt_1\right)$$
$$= \sum_{i>p}\sum_{j>p}\psi_i(t)\psi_j(t)\int\int K(t_1,s_1)\psi_i(t_1)\psi_j(s_1)ds_1 dt_1 .$$

The integrated expected squared error is then

$$\int E(e_p^2(t))dt = \sum_{i>p}\sum_{j>p}\left(\int \psi_i(t)\psi_j(t)dt\right)\int\int K(t_1,s_1)\psi_i(t_1)\psi_j(s_1)ds_1 dt_1$$
$$= \sum_{i>p}\int\int K(t_1,s_1)\psi_i(t_1)\psi_i(s_1)ds_1 dt_1 .$$

# Truncated KL expansion

Proof continued. We are interested in the optimization problem

$$\min \int E(e_p^2(t))dt\,, \quad \text{a.s.} \quad \int \psi_i(t)\psi_j(t)dt = \delta_{ij}\,.$$

We therefore minimize the objective function

$$Q = \sum_{i>p} \int \int K(t_1, s_1)\psi_i(t_1)\psi_i(s_1)ds_1 dt_1 - \lambda_i \left( \int \psi_i^2(t)dt - 1 \right)$$

by taking the following functional derivative and setting equal to 0:

$$\frac{dQ}{d\psi_i}(t) = 2\int K(t,s)\psi_i(s)ds - 2\lambda_i\psi_i(t) = 0\,.$$

Thus, we obtain the Fredholm equation

$$\int_{\mathcal{T}} K(t,s)\phi_i(s)ds = \lambda_i\phi_i(t)\,.$$

# Analytic example: Brownian motion on $\mathcal{T} = [0, 1]$

Let $s \leq t$, the covariance function for Brownian motion is

$$\begin{aligned} K(s,t) &= E(B_s B_t) = E\left(B_s(B_t - B_s + B_s)\right) \\ &= E(B_s^2) = s\,. \end{aligned}$$

Thus, the KL expansion corresponds to the integral equation

$$\int \min(s,t)\phi(s)ds = \lambda\phi(t)$$

or

$$\int_0^t s\phi(s)ds + t\int_t^1 \phi(s)ds = \lambda\phi(t)\,,$$

and thus $\phi(0) = 0$.

# Analytic example: Brownian motion on $\mathcal{T} = [0, 1]$

Taking the first derivative of both sides of

$$\int_0^t s\phi(s)ds + t\int_t^1 \phi(s)ds = \lambda\phi(t)$$

gives

$$\int_t^1 \phi(s)ds = \lambda\frac{d}{dt}\phi(t)\,.$$

Taking the second derivative gives the ODE

$$-\phi(t) = \lambda\frac{d^2}{dt^2}\phi(t)\,,$$

which is solved by

$$\phi(t) = A\sin(t/\sqrt{\lambda}) + B\cos(t/\sqrt{\lambda})\,.$$

# Analytic example: Brownian motion on $\mathcal{T} = [0, 1]$

... which is solved by

$$\phi(t) = A\sin(t/\sqrt{\lambda}) + B\cos(t/\sqrt{\lambda})\,.$$

But $\phi(0) = 0$, so $B = 0$ and $\phi(t) = A\sin(t/\sqrt{\lambda})$. To get eigenvalues, substitute $\phi(t)$ into the first derivative equations:

$$\begin{aligned}
A\int_t^1 sin(s/\sqrt{\lambda})ds &= A\lambda\cos(t/\sqrt{\lambda})/\sqrt{\lambda} && \Longrightarrow \\
A\sqrt{\lambda}\left(\cos(t/\sqrt{\lambda}) - \cos(1/\sqrt{\lambda})\right) &= A\sqrt{\lambda}\cos(t/\sqrt{\lambda}) && \Longrightarrow \\
\cos(1/\sqrt{\lambda}) &= 0 && \Longrightarrow \\
\lambda_i &= \frac{4}{(2i-1)^2\pi^2}\,, \quad i \geq 1\,.
\end{aligned}$$

# Analytic example: Brownian motion on $\mathcal{T} = [0, 1]$

Use the orthogonality of $\phi_i$ to get $A$:

$$\begin{aligned} 1 = \int_0^1 \phi_i^2(t)dt &= A^2 \int_0^1 \sin^2(t/\sqrt{\lambda})dt \\ &= A^2 \int_0^1 \sin^2\left(\left(i - \frac{1}{2}\right)\pi t\right) dt \\ &= A^2/2\,. \end{aligned}$$

Therefore, $A = \sqrt{2}$. Thus, the KL expansion is

$$B(t) = \sqrt{2} \sum_{i \geq 1} \frac{2\xi_i}{(2i-1)\pi} \sin\left(\left(i - \frac{1}{2}\right)\pi t\right), \quad \xi_i \overset{iid}{\sim} N(0,1)\,.$$

# Analytic example: Brownian bridge on $\mathcal{T} = [0, 1]$

Let B(t) be a standard Brownian motion and define the Brownian bridge $X(t)$ as

$$X(t) = B(t) - tB(1)\,.$$

The covariance function is $K(s, t) = \min(s, t) - st$. To see this, let $s \leq t$:

$$\begin{aligned} K(s,t) &= E(X(s)X(t)) = E((B_s - sB_1)(B_t - tB_1)) \\ &= E(B_s B_t) - sE(B_1 B_t) - tE(B_s B_1) + stE(B_1^2) \\ &= s - st - st + st = s - st\,. \end{aligned}$$

So the integral equation is

$$\int (\min(s,t) - st)\,\phi(s)ds = \lambda\phi(t)$$

# Analytic example: Brownian bridge on $\mathcal{T} = [0, 1]$

Differentiating the integral equation

$$\int_0^1 (\min(s, t) - st)\,\phi(s)ds = \lambda\phi(t)$$

w.r.t. $t$ gives

$$\int_t^1 \phi(s)ds - \int_0^1 s\phi(s) = \lambda \frac{d}{dt}\phi(t)\,,$$

and differentiating again gives

$$-\phi(t) = \lambda \frac{d^2}{dt^2}\phi(t)\,.$$

Just as with the standard Brownian motion, assuming $\phi(0) = 0$ results in $\phi(t) = A\sin(t/\sqrt{\lambda})$.

# Analytic example: Brownian bridge on $\mathcal{T} = [0, 1]$

Substituting $\phi(t) = A\sin(t/\sqrt{\lambda})$ into the first derivative equation

$$\int_t^1 \phi(s)ds - \int_0^1 s\phi(s) = \lambda \frac{d}{dt}\phi(t)$$

gives

$$\begin{aligned} A\int_t^1 sin(s/\sqrt{\lambda})ds - A\int_0^1 s\sin(s/\sqrt{\lambda})ds &= A\lambda\cos(t/\sqrt{\lambda})/\sqrt{\lambda} \\ \sqrt{\lambda}\left(\cos(t/\sqrt{\lambda}) - \cos(1/\sqrt{\lambda})\right) - & \\ \lambda\sin(1/\sqrt{\lambda}) + \sqrt{\lambda}\cos(1/\sqrt{\lambda}) &= \sqrt{\lambda}\cos(t/\sqrt{\lambda}) \\ \sin(1/\sqrt{\lambda}) &= 0 \\ \lambda_i &= \frac{1}{i^2\pi^2}\,, \quad i \geq 1\,. \end{aligned}$$

# Analytic example: Brownian bridge on $\mathcal{T} = [0, 1]$

Again, we use orthonormality of $\phi_i$ to solve for $A$:

$$1 = A^2 \int_0^1 \sin^2(s/\sqrt{\lambda_i}) ds = A^2 \int_0^1 \sin^2(i\pi s) ds = \frac{A^2}{2} \, .$$

Again, we have $A = \sqrt{2}$, and the KL expansion for the Brownian bridge may be written

$$\begin{aligned} X(t) &= \sum_{i \geq 1} \sqrt{\lambda_i} \phi_i(t) \xi_i \\ &= \sqrt{2} \sum_{i \geq 1} \frac{\xi_i}{\pi i} \sin(i\pi t) \, , \quad \xi_i \overset{iid}{\sim} N(0, 1) \, . \end{aligned}$$