STA547: HW2 *

Zhiling Gu

Contents

*Instructor: Xiongtao Dai, Iowa State University

1 Bivariate FPCA

The inner product of bivariate functions X^i . Let $X_1^i, X_2^i, i = 1, \ldots, n$ denote the first and second component of each observation X^i:

$$\hat{\mu}_1(t) = \frac{1}{n}\sum_{i=1}^{n} X_1^i(t)$$
$$\hat{\mu}_2(t) = \frac{1}{n}\sum_{i=1}^{n} X_2^i(t)$$
$$\hat{G}_{\alpha\beta}(t,s) = \frac{1}{n-1}\sum_{i=1}^{n}(X_\alpha^i(t) - \mu_\alpha(t))(X_\beta^i(s) - \mu_\beta(s)), \quad \alpha, \beta \in \{0, 1\}$$

Note that under the bivariate inner product defined, we need to solve the following system for the eigenvalues and eigenfunctions.

$$\int G_{12}(s,t)\xi_2(t)dt + \int G_{11}(s,t)\xi_1(t)dt = \lambda\xi_1(s)$$
$$\int G_{22}(s,t)\xi_2(t)dt + \int G_{21}(s,t)\xi_1(t)dt = \lambda\xi_2(s)$$

In the discrete approximation, we only need to solve

$$\sum_{j=1}^{m}\frac{(b-a)}{m}\hat{G}_{12}(s,t_j)\hat{\xi}_2(t_j) + \sum_{j=1}^{m}\frac{(b-a)}{m}\hat{G}_{11}(s,t_j)\hat{\xi}_1(t_j) = \lambda\xi_1(s)$$
$$\sum_{j=1}^{m}\frac{(b-a)}{m}\hat{G}_{22}(s,t_j)\hat{\xi}_2(t_j) + \sum_{j=1}^{m}\frac{(b-a)}{m}\hat{G}_{21}(s,t_j)\hat{\xi}_1(t_j) = \lambda\xi_2(s)$$

where $s \in t_1, \ldots, t_m$. Which lead us to solve

$$\frac{(b-a)}{m}\hat{G}_{12}\hat{\xi}_2 + \frac{(b-a)}{m}\hat{G}_{11}\hat{\xi}_1 = \lambda\xi_1$$
$$\frac{(b-a)}{m}\hat{G}_{22}\hat{\xi}_2 + \frac{(b-a)}{m}\hat{G}_{21}\hat{\xi}_1 = \lambda\xi_2$$
$$\Longleftrightarrow$$
$$\frac{(b-a)}{m}\begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \lambda\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}$$

Please see Figure 1 for the results illustration.□

2 Simulation of Gaussian and Non-Gaussian Processes

Recall that if $X(t), t \in \mathcal{T}$ is a L^2 stochastic process, then (i) ξ_i are independent and $\xi_k \sim N(0, \lambda_k)$. (ii)

$$\begin{pmatrix} \xi_k \\ X(t) \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ \mu(t) \end{pmatrix}, \begin{pmatrix} \lambda_k & \lambda_k\phi_k'(t) \\ \lambda_k\phi_k(t) & G(t,t) \end{pmatrix}\right) \tag{1}$$

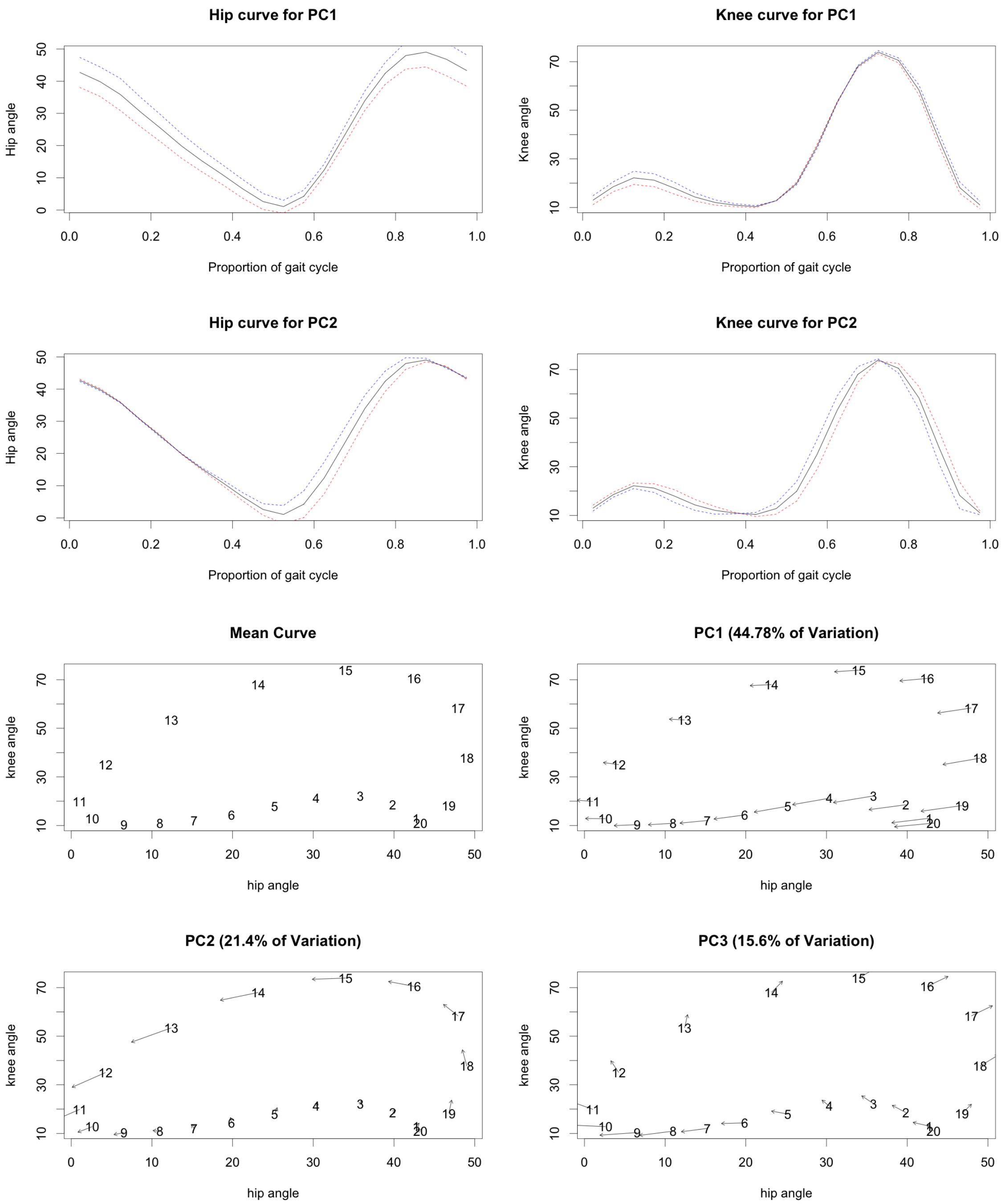
To simulate Gaussian Process, we need to set ϕ_k, $\lambda_k, k = 1, \ldots, K$ and simulate $\xi_{ik}, i = 1, \ldots, n, k = 1, \ldots, K$. Algorithm is as follows

1. $\xi_i = Z_i' Diag(\lambda_1, \ldots, \lambda_K)$ where $Z_i \sim N(0_K, Diag(1_K))$
2. $X_i(t) = \mu'(t) + \xi_i\phi'(t)$ where columns of $\phi(t)$ is the PCs

In FDA, it is better to use $\top$ for transpose

To simulate Non-Gaussian Process, recall the properties of FPCs (i) $E(\xi_k) = 0$, (ii) $Cov(\xi_k, \xi_{k'}) = \lambda_k\xi_{kk'}$. Let $\xi_k \sim U(-a_k, a_k)$ with $Var(\xi_k) = a_k^2/3 \equiv \lambda_k \implies a_k = \sqrt{3\lambda_k}$ and $E(\xi_k) = 0$. The algorithm is as follows

Figure 1: Bivariate FPCA



1. $\xi_i = (u_1, \ldots, u_K)$ where $u_i \sim U(-\sqrt{3\lambda_i}, \sqrt{3\lambda_i}) \iff \xi_i = (u_1, \ldots, u_K)Diag(\sqrt{3\lambda_1}, \ldots, \sqrt{3\lambda_K}), u_i \sim U(-1, 1)$
2. $X_i(t) = \mu'(t) + \xi_i \phi'(t)$ where columns of $\phi(t)$ is the PCs

Normality tests are also performed to see of the linear combination of $X(t)$ is still normal. In our case, we tested the normality of $\bar{X}_i = 1/m \sum_{j=1}^{m} X_i(t_j), i = 1, \ldots, 200$. The gaussian process produces $\bar{X}_i$ with distribution similar normal, while the non-gaussian process does not. Please see Figure 2 for the results illustration.□

3 Implement FPCA

Implement an FPCA for the yeast data. Describe the first few modes of variation, and discuss whether the variation in the dataset is properly summarized by the first few eigenfunctions using FPCA.

The yeast data contains missing data. The approach applied here is to first omit the whole observation of a time when there are observations missing for certain genes. Then do linear splines on mean and covariance matrix

One could use the ~~omit~~ data available at each time pt to estimate the mean & cov, without discarding the time pt

Figure 2: Stochastic Process Simulation

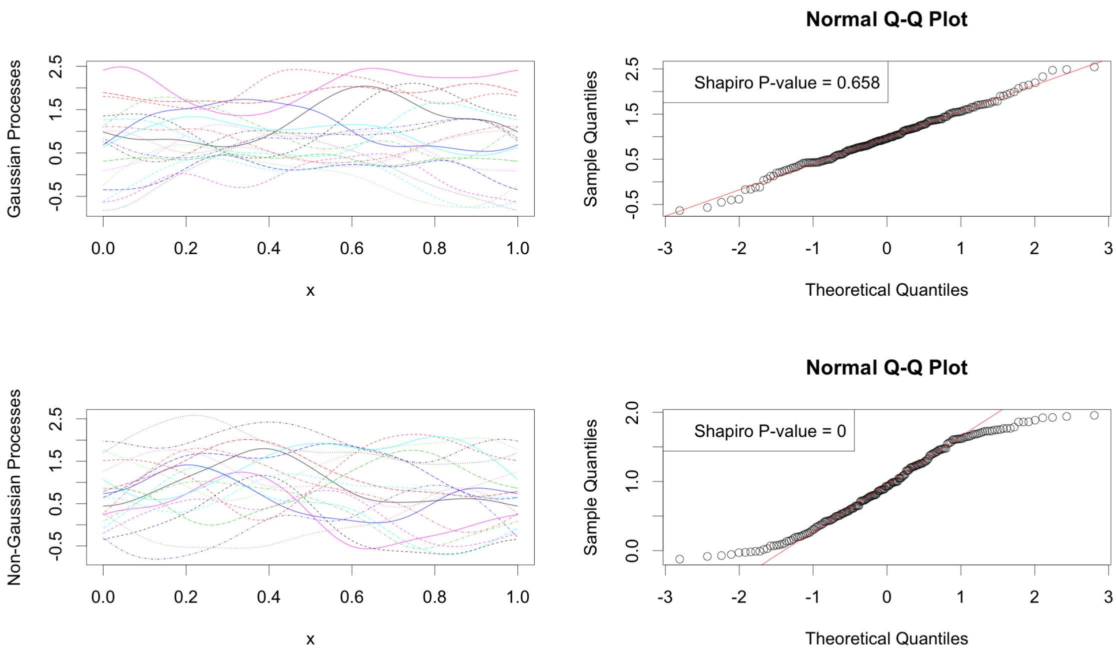
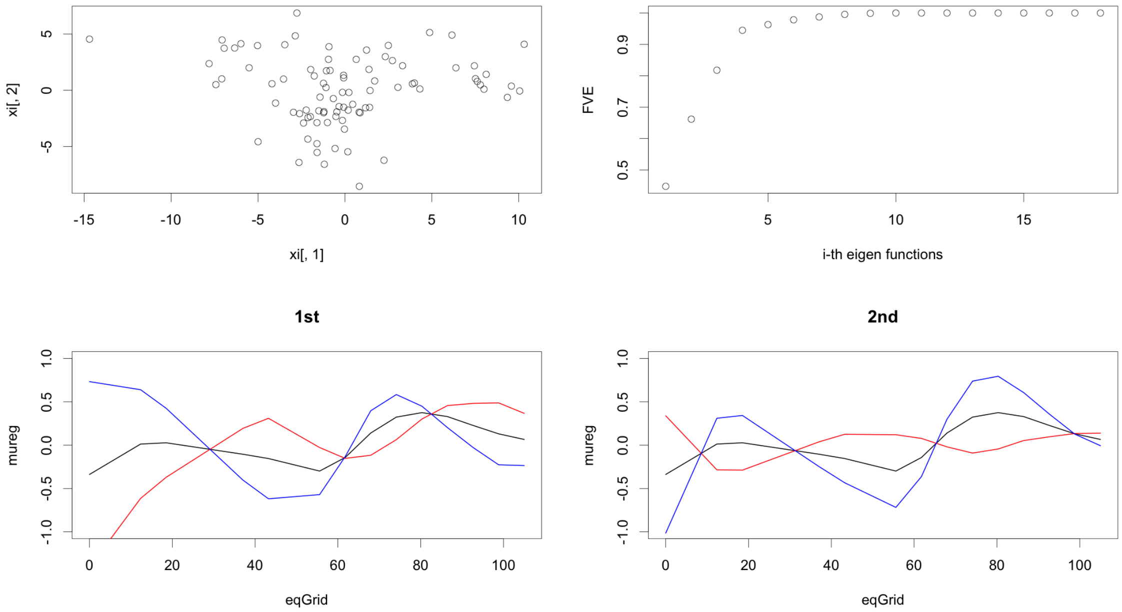


Figure 3: Yeast data analysis



to get regular grid FPCs. From the plot of fraction of variation (FVE), we are able to see the first 4 eigen functions explained over 90% of the variations. Please for results illustration. $\square$

4 Properties of eigen functions Estimates

Let $X_1, \ldots, X_n$ be independent realizations of an L^2 stochastic process X on $\mathcal{T} = [0,1]$. Let $\left(\hat{\lambda}_k, \hat{\phi}_{ik}\right)$ be the k th eigenvalue eigenfunction pair of the sample covariance function $\hat{G}$. Let $Z_k = n^{-1} \sum_{i=1}^{n} \hat{\xi}_{ik}$. Show that

(a) $Z_k = 0$ for $k = 1, \ldots, n-1$;

$$
\begin{aligned}
Z_k = \frac{1}{n}\sum_{i=1}^{n}\hat{\xi}_{ik} &= \frac{1}{n}\sum_{i=1}^{n}\int_{\mathcal{T}}(X_i(t)-\hat{\mu}(t))\hat{\phi}(t)dt \\
&= \frac{1}{n}\int_{\mathcal{T}}\left[\sum_{i=1}^{n}X_i(t)-n\mu(t)\right]\hat{\phi}_k(t)dt \\
&= \frac{1}{n}\int_{\mathcal{T}}[n\mu(t)-n\mu(t)]\hat{\phi}_k(t)dt = 0\,\square
\end{aligned}
$$

(b) $(n-1)^{-1}\sum_{i=1}^{n}\left(\hat{\xi}_{ik}-Z_k\right)^2 = \hat{\lambda}_k$

$$
\begin{aligned}
\frac{1}{n-1}\left(\sum_{i=1}^{n}\hat{\xi}_{ik}-Z_k\right)^2 &\overset{(a)}{=} \frac{1}{n-1}\sum_{i=1}^{n}\hat{\xi}_{ik}^2 \\
&= \frac{1}{n-1}\int_{\mathcal{T}}\int_{\mathcal{T}}(X_i(t)-\hat{\mu}(t))\hat{\phi}_k(t)(X_i(s)-\hat{\mu}(s))\hat{\phi}_k(s)dtds \\
&= \frac{1}{n-1}\int_{\mathcal{T}}\int_{\mathcal{T}}(n-1)\hat{G}(t,s)\hat{\phi}_k(t)\hat{\phi}_k(s)dtds \\
&= \int_{\mathcal{T}}\hat{\lambda}_k\hat{\phi}_k(s)\hat{\phi}_k(s)ds \\
&= \hat{\lambda}_k\,\square
\end{aligned}
$$

The last equality holds since $\hat{\phi}_k$ is orthonomal eigen function and has norm 1.

(c) $(n-1)^{-1}\sum_{i=1}^{n}\left(\hat{\xi}_{ik}-Z_k\right)\left(\hat{\xi}_{ik'}-Z_{k'}\right) = 0$ if $k \neq k'$

$$
\begin{aligned}
\frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{\xi}_{ik}-Z_k\right)\left(\hat{\xi}_{ik'}-Z_{k'}\right) &= \frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{\xi}_{ik}\hat{\xi}_{ik'}-Z_k\hat{\xi}_{ik'}-Z_{k'}\hat{\xi}_{ik}+Z_kZ_{k'}\right) \\
&\overset{(a,b)}{=} \frac{1}{n-1}\left(\sum_{i=1}^{n}\hat{\xi}_{ik}\hat{\xi}_{ik'}\right) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\int_{\mathcal{T}}\int_{\mathcal{T}}(X_i(t)-\hat{\mu}(t))\hat{\phi}_k(t)(X_i(s)-\hat{\mu}(s))\hat{\phi}_{k'}(s)dtds \\
&= \frac{1}{n-1}\int_{\mathcal{T}}\int_{\mathcal{T}}\sum_{i=1}^{n}(X_i(t)-\hat{\mu}(t))(X_i(s)-\hat{\mu}(s))\hat{\phi}_k(t)dt\hat{\phi}_{k'}(s)ds \\
&= \frac{1}{n-1}\int_{\mathcal{T}}\int_{\mathcal{T}}(n-1)\hat{G}(s,t)\hat{\phi}_k(t)dt\hat{\phi}_{k'}(s)ds \\
&= \int_{\mathcal{T}}\hat{\lambda}_k\hat{\phi}_k(s)\hat{\phi}_{k'}(s)ds \\
&= \hat{\lambda}_k\int_{\mathcal{T}}\hat{\phi}_k(s)\hat{\phi}_{k'}(s)ds = 0\,\square
\end{aligned}
$$

5 Tensor Product Operator

Let $x_1,\ldots,x_n$ be n linearly independent (nonrandom) functions in a Hilbert space H. Show that $\sum_{i=1}^{n}(x_i-\overline{x})\otimes(x_i-\overline{x}) \in \mathcal{B}(\mathbb{H})$ has rank $n-1$

Consider $y_i := x_i - \bar{x}$. Since $x_i \in \mathbb{H}, \forall i$, $y_i \in \mathbb{H}$. Let $\mathcal{T} := \sum_{i=1}^{n}(x_i-\bar{x})\otimes(x_i-\bar{x}) = \sum_{i=1}^{n}y_i\otimes y_i$ denote the tensor product operator. Let $\Im(\mathcal{T}) := \{\sum_{i=1}^{n}\langle y_i, y\rangle y_i : y \in \mathbb{H}\}$ denote the image of the operator from $\mathbb{H}$ to $\mathbb{H}$. Note $\sum_{i=1}^{n}y_i = 0$, $\dim(\mathrm{Span}\{y_1,\ldots,y_n\}) = n-1$

Now consider an complete orthonormal basis $e_1,\ldots,e_{n-1}$ of the smallest Hilbert space $\mathbb{H}_0$ that contains $y_1,\ldots,y_n$ where $y_i = \sum_{k=1}^{n-1}c_{ik}e_k$, $c_{ik} = \langle y_i, e_k\rangle$, $\mathbb{H}_0 \subset \mathbb{H}$. Let $\Im_0(\mathcal{T}) := \{\sum_{i=1}^{n}\sum_{k=1}^{n-1}c_{ik}\langle e_k, y\rangle y_i, y \in \mathbb{H}_0\}$.

Obviously $\Im_0(\mathcal{T}) \subset \Im(\mathcal{T})$.

Note y is any element in $\mathbb{H}_0 \implies \langle e_k, y\rangle_{k=1,\ldots,n-1}$ spans $\mathbb{R}^{n-1} \implies \sum_{k=1}^{n-1} c_{ik}\langle e_k, y\rangle$ spans $\mathbb{R}$. Therefore $\Im_0(\mathcal{T}) = \text{Span}\{y_1, \ldots, y_n\}$.

Also $\Im(\mathcal{T}) \subset \text{Span}\{y_1, \ldots, y_n\}$ by definition. Therefore $\Im(\mathcal{T}) = \text{Span}\{y_1, \ldots, y_n\}$, $\dim(\Im(\mathcal{T})) = \dim(\Im_0(\mathcal{T})) = n-1$. Therefore the rank of $\mathcal{T}$ is the dimension of its image, equal to $(n-1)$. $\square$

6 (Reproducing Kernel Hilbert Space)

Let $K : [0,1] \times [0,1] \to \mathbb{R}$ be a continuous symmetric non-negative positive definite function (Mercer's kernel). Define

$$\mathcal{H}_K = \left\{ f = \sum_{j=1}^{\infty} a_j \lambda_j e_j | \sum_{j=1}^{\infty} a_j^2 \lambda_j < \infty \right\} \subset L^2$$

where the (λ_j, e_j) are the eigenvalue and eigenfunction pairs of the integral operator $\mathscr{K}$ associated with K. For $f = \sum_{j=1}^{\infty} a_j \lambda_j e_j$ and $g = \sum_{j=1}^{\infty} b_j \lambda_j e_j$ in $\mathcal{H}_K$, define inner product

$$\langle f, g\rangle_K = \sum_{j=1}^{\infty} a_j b_j \lambda_j$$

Then $\mathcal{H}_K$ is a Hilbert space with this norm.

(a) Show that $f(x) = \sum_{j=1}^{\infty} a_j \lambda_j e_j(x)$, where the sum converges absolutely, and uniformly over $x \in [0,1]$, i.e. $\sup_x \sum_{j=n}^{\infty} |a_j \lambda_j e_j(x)| \to 0$ as $n \to \infty$.
(b) Show that $K(\cdot, t) \in \mathcal{H}_K$ for $t \in [0,1]$. (c) Show that $\langle K(\cdot,t), f\rangle_K = f(t)$ for $t \in [0,1]$. This means the evaluation functional $\delta_t : \mathcal{H}_K \to \mathbb{R}, \delta_t(f) = f(t)$ is a continuous linear functional. (Remark: A Hilbert space of functions in which point evaluation is a continuous linear functional is called a Reproducing Kernel Hilbert Space (RKHS).)

Proof:

Firstly prove (c). Recall the definition of integral operator:

$$\mathscr{K}(f)(\cdot) = \int K(\cdot, s) f(s) ds \tag{2}$$

Since K is a symmetric non-negative definite function, the eigen values of $\mathscr{K}$, i.e. λ_j, should all be non-negative. By Mercer's theorem, K has the representation:

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_j e_j(s) e_j(t) \tag{3}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ are the eigen values and $e_1, e_2, \ldots$ are the orthonormal eigen functions of $\mathscr{K}$. And the sum above converges absolutely and uniformly.

Then

$$\begin{aligned}\langle K(\cdot,t), f\rangle_K &= \langle \sum_{j=1}^{\infty} \lambda_j e_j(t) e_j, \sum_{j=1}^{\infty} a_j \lambda_j e_j\rangle_K \\ &= \sum_{j=1}^{\infty} \lambda_j e_j(t) a_j \\ &= f(t)\end{aligned}$$

Since $f \in L^2[0,1]$, $f(t)$ must be bounded and therefore the equation holds. $\square$

Secondly, prove (b). We have recalled from Mercer's theorem that,

$$K(\cdot, t) = \sum_{j=1}^{\infty} \lambda_j e_j(t) e_j$$

Check whether $a_j = e_j(t)$ and λ_j satisfies $\sum_{j=1}^{\infty} a_j^2 \lambda_j < \infty$:

$$\sum_{j=1}^{\infty} a_j^2 \lambda_j = \sum_{j=1}^{\infty} e_j^2(t) \lambda_j = K(t,t) < \infty$$

which is finite because K is continuous on a bounded closed region and thus must be finite.

Finally prove (a). Refer to HE Theorem 2.7.6: Convergence in Hilbert norm implies pointwise (uniform) convergence. Let $f_n := \sum_{j=1}^{n} a_j \lambda_j e_j$. Since $f = \sum_{j=1}^{\infty} a_j \lambda_j e_j \in \mathcal{H}_K$. By reproducing property (c) and the definition of RKHS, we have

$$\begin{aligned} \|f - f_n\|^2 &= \langle f - f_n, f - f_n \rangle_K \\ &= \langle \sum_{j>n} a_j \lambda_j e_j, \sum_{j>n} a_j \lambda_j e_j \rangle_K \\ &= \sum_{j>n} a_j^2 \lambda_j \to 0 \text{ as } n \to \infty \end{aligned}$$

Then

$$\begin{aligned} |f(t) - f_n(t)|^2 &= \langle f - f_n, K(\cdot, t) \rangle_K \\ &\overset{CS}{\leq} \|f - f_n\|^2 \langle K(\cdot, t), K(\cdot, t) \rangle_K \\ &= \|f - f_n\|^2 K(t,t) \\ &\to 0 \text{ as } n \to \infty \text{ since } K(\cdot, \cdot) < \infty \text{ by (b)} \ \square \end{aligned}$$

A