MA 540 — Project 1
February 8, 2018

Problem 1

Statement

Consider the covariance function $C(x,y) = \min(x,y)$ on the interval $[0,1]$ along with the mean function $\bar{\alpha}(x) = \frac{1}{8}(x+1)^2$. Compute the first $N = 3$ eigenvalues and eigenfunctions and plot the eigenfunctions. Now compute and plot 1000 realizations of the Karhunen-Loeve expansion truncated at $N = 3$ with random variables $Q_n \sim N(0,1)$. You can compute the random variables using the command `randn.m`.

For your 1000 realizations, use `histnorm.m` to plot a histogram scaled to unity of the realizations of $\beta(\bar{x},\omega) = \sum_{n=1}^{3} \sqrt{\lambda_n}\phi_n(\bar{x})Q_n(\omega)$ at $\bar{x} = 0.5$. On the same figure, plot a kernel density estimate (KDE) of the distribution and a normal approximation to it. Is the distribution unbiased and consistent with your choice of $\sigma^2 = 1$? Finally, repeat this plot with $N = 6$ to see if your expansion has converged.

Note: Use the property that

$$\operatorname{var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \operatorname{var}(X_i)\,.$$

Solution

Using the typeset lecture notes, we have eigenvalues and eigenfunctions

$$\lambda_n = \frac{1}{\pi^2(n-\frac{1}{2})^2}, \quad \phi_n(x) = \sqrt{2}\sin\left(\frac{x}{\sqrt{\lambda_n}}\right), \qquad n = 1,2,3,\ldots$$

The first three pairs are

$$\begin{aligned}
\lambda_1 &= \frac{4}{\pi^2}, \quad \phi_1(x) = \sqrt{2}\sin\left(\frac{\pi}{2}x\right), \\
\lambda_2 &= \frac{4}{9\pi^2}, \quad \phi_2(x) = \sqrt{2}\sin\left(\frac{3\pi}{2}x\right), \\
\lambda_3 &= \frac{4}{25\pi^2}, \quad \phi_3(x) = \sqrt{2}\sin\left(\frac{5\pi}{2}x\right).
\end{aligned}$$

Figure 1 shows plots of these eigenfunctions.

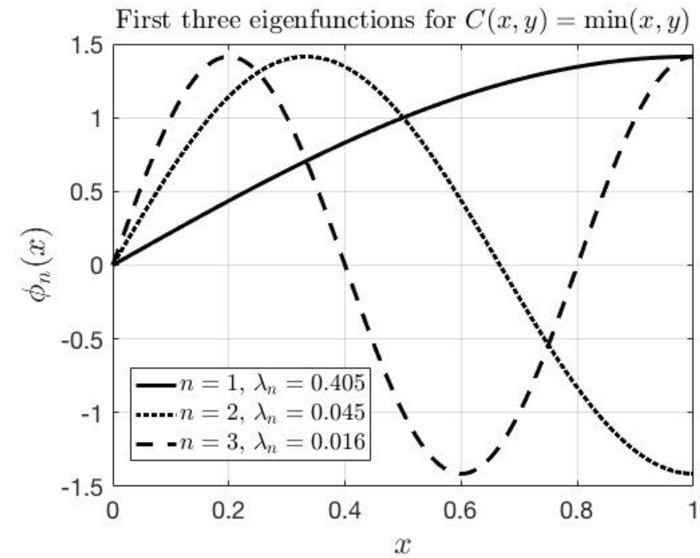

Figure 1: Eigenvalues and eigenfunctions for $C(x, y) = \min(x, y)$.

Using the using the `randn` command within a `for`-loop, we compute 1000 realizations of the Karhunen–Loève expansion truncated at $N = 3$:

$$\alpha(x, \omega) = \bar{\alpha}(x, \omega) + \sum_{n=1}^{3} \sqrt{\lambda_n}\phi_n(x)Q_n(\omega),$$

where $Q_n \sim N(0, 1)$. Figure 2 displays these 1000 realizations along with the sample mean. Note that with 1000 realizations, the sample mean closely follows $\bar{\alpha}(x)$, as we would hope.

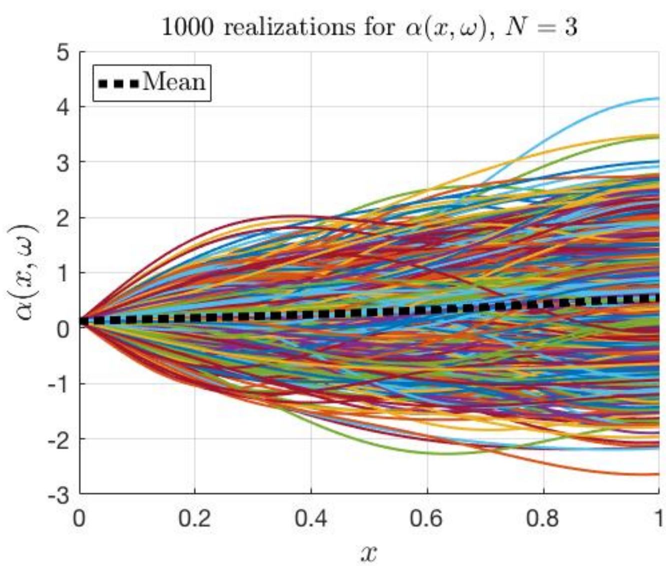

Figure 2: 1,000 realizations of the Karhunen–Loève expansion. Sample mean is shown for reference.

At each realization, we stored the value of $\beta(0.5, \omega) = \sum_{n=1}^{3} \sqrt{\lambda_n}\phi_n(0.5)Q_n(\omega)$. Figure 3a displays a histogram, kernel density estimate, and approximating normal of the 1000 realizations of $\beta(0.5, \omega)$.

From

$$\operatorname{var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 \operatorname{var}(X_i),$$

we get $\operatorname{var}(\beta(0.5,\omega)) = \sum_{n=1}^{3} \lambda_n \phi_n^2(0.5)\operatorname{var}(Q_n) = \sum_{n=1}^{3} \lambda_n \phi_n^2(0.5) \approx 0.4665$. So $\sigma = 0.683$ in our approximating normal distribution, which very closely matches the KDE. The sample mean of the histogram has magnitude on the order of $10^{-2} \approx 0$, so the distribution is unbiased.

Now we perform 1,000 realizations of $\beta(0.5,\omega)$ with $N = 6$ terms in the Karhunen–Loève expansion. Figure 3b shows the resulting histogram, KDE, and approximating normal. In the new plot, the KDE still closely tracks the approximating normal distribution, the sample mean still has magnitude on the order of 10^{-2}, and σ has changed only slightly (from 0.683 to 0.695). Thus, we conclude that our expansion has converged.

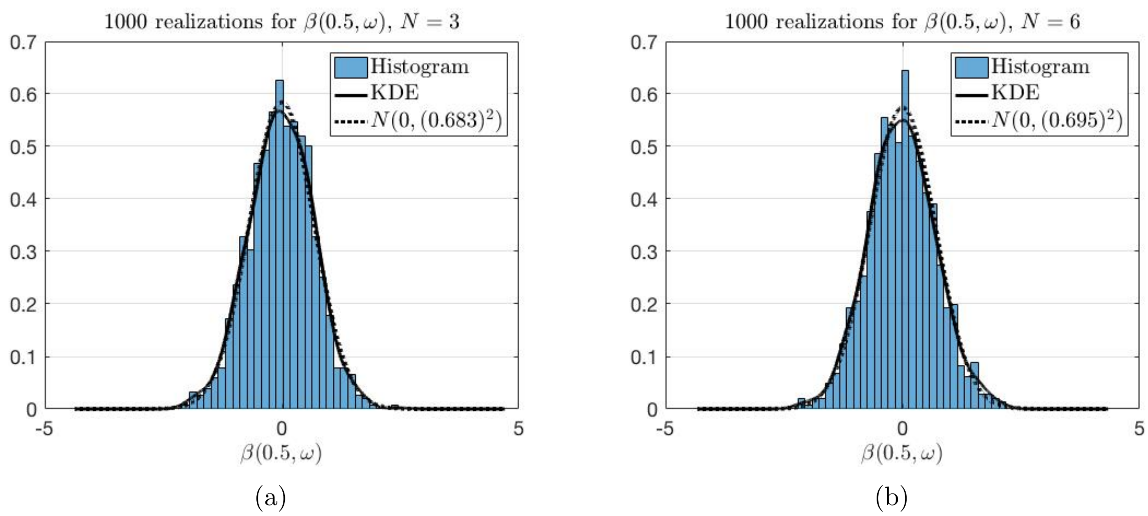

Figure 4: Histogram, KDE, and approximating normal for 1000 realizations of $\beta(0.5,\omega)$ with (a) $N = 3$ and (b) $N = 6$ terms in the KL expansion.

Problem 2

Statement

For the same mean function $\bar{\alpha}(x) = \frac{1}{8}(x+1)^2$, now consider the exponential covariance function $C(x,y) = e^{-|x-y|/L}$ on the interval $[-1,1]$ with the correlation length $L = 5$. The eigenvalues are

$$\lambda_n = \begin{cases} \frac{2L}{1+L^2 w_n^2} & , \text{ if } n \text{ is even,} \\ \frac{2L}{1+L^2 v_n^2} & , \text{ if } n \text{ is odd,} \end{cases}$$

and the eigenfunctions are

$$\phi_n(x) = \begin{cases} \frac{\sin(w_n x)}{\sqrt{1-\frac{\sin(2w_n)}{2w_n}}} & , \text{ if } n \text{ is even,} \\ \frac{\cos(v_n x)}{\sqrt{1+\frac{\sin(2v_n)}{2v_n}}} & , \text{ if } n \text{ is odd.} \end{cases}$$

Here w_n and v_n are solutions of the transcendental equations

$$\begin{cases} Lw + \tan(w) = 0 & , \quad \text{for even } n, \\ 1 - Lv\tan(v) = 0 & , \quad \text{for odd } n. \end{cases}$$

Compute the first odd and even roots v and w. You can get good initial guesses by plotting the functions and zooming to approximate the roots. You can obtain accurate values using the MATLAB command `fzero.m`. Compute the first two eigenvalues and eigenfunctions and plot the eigenfunctions. Now compute and plot 1000 realizations of the Karhunen-Loeve expansion truncated at $N = 2$ with random variables $Q_n \sim N(0,1)$.

Solution

By inspection of the plots for $f(w) = Lw + \tan(w)$ and $g(v) = 1 - Lv\tan(v)$, we arrive at the initial guesses

$$v_1^{(0)} = 0.43, \quad w_2^{(0)} = 1.68.$$

Using `fzero` with these initial iterates and a function-value tolerance of 10^{-8}, we arrive at

$$v_1 = 0.4328, \quad w_2 = 1.6887.$$

These values yield

$$\lambda_1 = \frac{2(5)}{1 + (5^2)(0.4328)^2} \approx 1.7594, \qquad \phi_1(x) = \frac{\cos(0.4328x)}{\sqrt{1 + \frac{\sin(2(0.4328))}{2(0.4328)}}},$$

$$\lambda_2 = \frac{2(5)}{1 + (5^2)(1.6887)^2} \approx 0.1383, \qquad \phi_2(x) = \frac{\sin(1.6887x)}{\sqrt{1 - \frac{\sin(2(1.6887))}{2(1.6887)}}}.$$

Figure 5 shows a plot of $\phi_1(x)$ and $\phi_2(x)$.

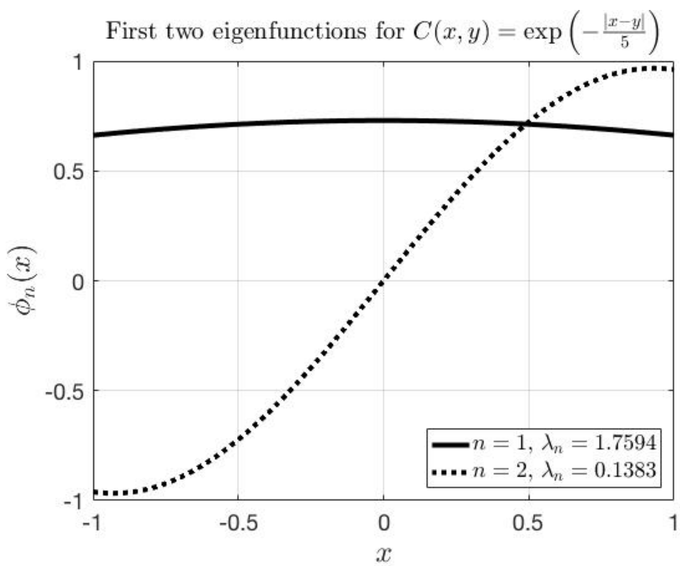

Figure 5: First two eigenfunctions for $C(x,y) = \exp\left(-\frac{|x-y|}{5}\right)$.

Now we perform 1000 realizations of the KL expansion truncated at $N = 2$ terms:

$$\alpha(x, \omega) = \bar{\alpha}(x) + \sum_{n=1}^{2} \sqrt{\lambda_n} \phi_n(x) Q_n(\omega), \qquad Q_n \sim N(0, 1) \,.$$

Figure 6 shows these realizations along with the sample mean, which closely tracks the true $\bar{\alpha}(x)$.

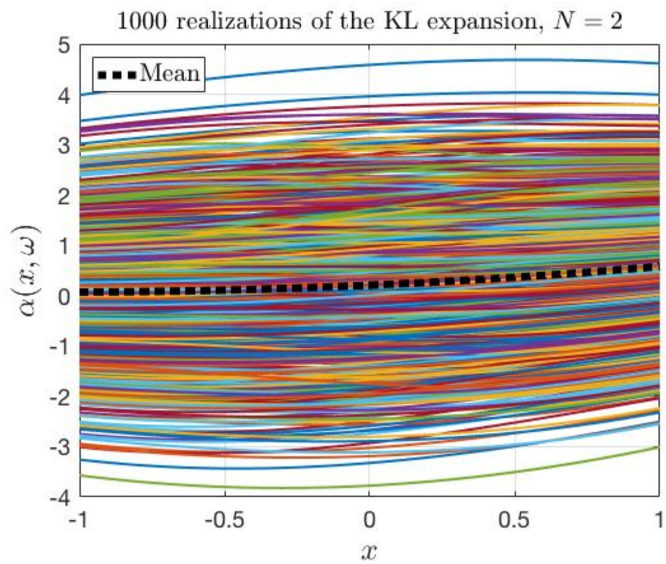

Figure 6: 1000 realizations of the Karhunen–Loève expansion truncated at $N = 2$ terms. Sample mean is shown for reference.

Problem 3

Statement

Consider again the covariance function $C(x, y) = \min(x, y)$ on the interval $[0, 1]$ and a Karhunen-Loeve expansion truncated at $N = 3$. You can ignore the contributions from the random variables Q_n. Your objective is to approximate the mean $\bar{\alpha}(x)$ using the representation $\bar{\alpha}(x, q) = q_0 + q_1 x + q_2 x^2$, where $q = [q_0, q_1, q_2]$ are parameters to be estimated using the data in `KL_data.txt`. In the file, the first column is the x-values at the 21 points $x_j = 0.05(j - 1)$ and the second column is values of

$$\begin{aligned} y_j &= \alpha(x_j, q) + \varepsilon_j \\ &= \bar{\alpha}(x_j, q) + \sum_{n=1}^{3} \sqrt{\lambda_n} \phi_n(x_j) + \varepsilon_j \end{aligned}$$

where ε_j is added measurement noise.

You should use optimization software to solve the minimization problem

$$q_{opt} = \arg\min_q \sum_{j=1}^{21} [y_j - \alpha(x_j, q)]^2 \,.$$

Plot your approximation to $\alpha(x, q)$ along with the data.

Solution

Following the email to the class, we use the eigenpairs

$$\lambda_n = \frac{1}{\pi^2(n+\frac{1}{2})^2}, \quad \phi_n(x) = \sqrt{2}\sin\left(\frac{x}{\sqrt{\lambda_n}}\right), \qquad n = 1, 2, 3, \ldots$$

in the data fitting. (Subtraction has changed to addition in the denominator of λ.) We truncate the KL expansion to include only the first three terms given by the relationship above. For specificity, the eigenvalues and eigenfunctions we use are

$$\begin{aligned}
\lambda_1 &= \frac{4}{9\pi^2}, \quad \phi_1(x) = \sqrt{2}\sin\left(\frac{3\pi}{2}x\right), \\
\lambda_2 &= \frac{4}{25\pi^2}, \quad \phi_2(x) = \sqrt{2}\sin\left(\frac{5\pi}{2}x\right), \\
\lambda_3 &= \frac{4}{49\pi^2}, \quad \phi_3(x) = \sqrt{2}\sin\left(\frac{7\pi}{2}x\right).
\end{aligned}$$

Our goal is to find the minimizer of

$$J(\mathbf{q}) = \sum_{j=1}^{21}\left[q_0 + q_1 x_j + q_2 x_j^2 + \left(\sum_{n=1}^{3}\sqrt{\lambda_n}\phi_n(x_j)\right) - y_j\right]^2.$$

Using `fminsearch` with initial iterate $\mathbf{q}^0 = \mathbf{0}$ and function-value and x-value tolerances set to 10^{-8}, we obtain the optimal parameter values

$$\mathbf{q}_{opt} = [q_0^{opt}, q_1^{opt}, q_2^{opt}] = [0.2002, -0.7968, 0.7973]$$

and optimal cost $J(\mathbf{q}_{opt}) = 0.0033$. Figure 7 displays the data along with the best-fit curve, given by

$$\alpha(x;\mathbf{q}_{opt}) = 0.2002 - 0.7968x + 0.7973x^2 + \sum_{n=1}^{3}\sqrt{\lambda_n}\phi_n(x).$$

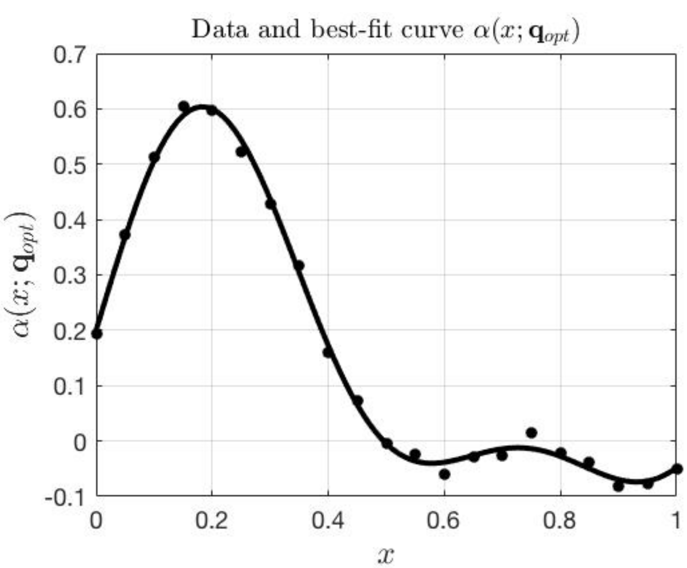

Figure 7: Data and $\alpha(x;\mathbf{q}_{opt})$.