Jared Yu

Computational Statistics

Problem Set 10

NOTE: The code for the assignment is in the Appendix.

1. Show that the Fourier Trigonometric Family

$$\{1, \cos x, \sin x, \cos 2x, \sin 2x, \cos 3x, \sin 3x, \cdots\}$$

is an orthogonal family over $0 \leq x \leq 2\pi$ with respect to the weight function $w(x) = 1$.

The Fourier Trigonometric Family, denoted $\{q_i(x)\}$, is orthogonal over $D = 0 \leq x \leq 2\pi$ w.r.t. $w(x) = 1$ i.f.f.

$$\langle q_i, q_j \rangle = \int_D q_i(x) q_j(x) w(x)\, dx = \begin{cases} 0 & i \neq j \\ \lambda_i > 0 & i = j \end{cases}.$$

To show this it will first be noted that there are three possible cases of $\langle q_i, q_j \rangle$ based combinations of the sine and cosine functions within $\{q_i(x)\}$. Next, it will be noted that the first element of $\{q_i(x)\}$, $q_1(x) = 1$, can be represented with $\cos(ix)$ for $i = 0$ since $\cos(0) = 1$. Then all the pairwise combinations of members of $\{q_i(x)\}$ can be denoted as follows:

$$\int_D \sin(ix)\sin(jx)\, dx = \begin{cases} 0 & i \neq j; i,j > 0 \\ \pi & i = j; i,j > 0 \end{cases} \tag{1}$$

$$\int_D \cos(ix)\cos(jx)\, dx = \begin{cases} 2\pi & i = j = 0 \\ 0 & i \neq j; i,j > 0 \\ \pi & i = j; i,j > 0 \end{cases} \tag{2}$$

$$\int_D \sin(ix)\cos(jx)\, dx = 0 \quad i > 0, j \geq 0 \tag{3}$$

Below, each of these equations and their cases will be shown:

Equation (1)

Case 1: $i \neq j$, $i,j > 0$

$$\int_D \sin(ix)\sin(jx)\, dx = \frac{1}{2}\int_D \cos\big((i-j)x\big) - \cos\big((i+j)x\big)\, dx$$

$$= \frac{1}{2}\left\{\frac{1}{i-j}\sin\big((i-j)x\big)\,\Big|_0^{2\pi} - \frac{1}{i+j}\sin\big((i+j)x\big)\Big|_0^{2\pi}\right\}$$

$$= \frac{1}{2}\left\{\left[\frac{1}{i-j}\Big(\sin\big((i-j)2\pi\big) - \sin\big((i-j)0\big)\Big)\right] - \left[\frac{1}{i+j}\Big(\sin\big((i+j)2\pi\big) - \sin\big((i+j)0\big)\Big)\right]\right\} = 0$$

Case 2: $i = j$; $i,j > 0$

$$\int_D \sin(ix)\sin(jx)\,dx = \frac{1}{2}\int_D \cos(0) - \cos(2ix)\,dx$$

$$= \frac{1}{2}\left\{\int_D 1 - \cos(2ix)\,dx\right\} = \frac{1}{2}\left\{2\pi - \frac{1}{2i}\sin(2ix)\,|_0^{2\pi}\right\} = \frac{1}{2}\{2\pi - 0\} = \pi$$

Equation (2)

Case 1: $i = j = 0$

$$\int_D \cos(0)\cos(0)\,dx = \int_D 1\,dx = x|_0^{2\pi} = 2\pi$$

Case 2: $i \neq j;\ i, j > 0$

$$\int_D \cos(ix)\cos(jx)\,dx = \frac{1}{2}\int_D \cos\big((j+k)x\big) + \cos\big((j-k)x\big)\,dx$$

$$= \frac{1}{2}\left[\frac{1}{j+k}\sin\big((j+k)x\big)\,|_0^{2\pi} + \frac{1}{j-k}\sin\big((j-k)x\big)|_0^{2\pi}\right]$$

$$= \frac{1}{2}\left\{\left[\frac{1}{j+k}\Big(\sin\big((j+k)2\pi\big) - \sin(0)\Big)\right] + \left[\frac{1}{j-k}\Big(\sin\big((j-k)2\pi\big) - \sin(0)\Big)\right]\right\} = 0$$

Case 3: $i = j;\ i, j > 0$

$$\int_D \cos(ix)\cos(jx)\,dx = \frac{1}{2}\int_D \cos(2ix) + 1\,dx$$

$$= \frac{1}{2}\left[\frac{1}{2i}\sin(2ix)\,|_0^{2\pi} + 2\pi\right] = \pi$$

Equation (3)

Case 1: $i > 0, j \geq 0$

$$\int_D \cos(ix)\sin(jx)\,dx = \frac{1}{2}\left\{\int_D \sin\big((i+j)x\big) + \sin\big((i-j)x\big)\,dx\right\}$$

$$= \frac{1}{2}\left\{-\frac{1}{i+j}\cos\big((i+j)x\big)|_0^{2\pi} + \left(-\frac{1}{i-j}\right)\cos\big((i-j)x\big)|_0^{2\pi}\right\}$$

$$= \frac{1}{2}\left\{\left[\left(-\frac{1}{i+j}\right)\Big(\cos\big((i+j)2\pi\big) - \cos(0)\Big)\right] + \left[\left(-\frac{1}{i-j}\right)\Big(\cos\big((i-j)2\pi\big) - \cos(0)\Big)\right]\right\} = 0$$

In the above 3 equations, the property of $\langle q_i, q_j\rangle = \begin{cases} 0 & i \neq j \\ \lambda_i > 0 & i = j \end{cases}$ holds and so it could be said that the Fourier Trigonometric Family is orthogonal. ■

2. (a) Show that if $\{q_i(x)\}$ is a set of orthogonal functions, then it is a linearly independent set.

A set $\{q_i(x)\}$ is linearly independent if

$$\sum_i c_i q_i(x) = 0$$

only when $c_i = 0 \; \forall \; i$.

To show that if $\{q_i(x)\}$ is a set of orthogonal function, then it is a linearly independent set, suppose then that $\{q_i(x)\}$ isn't linearly independent. Then for some $c_i \neq 0$,

$$\sum_i c_i q_i(x) = 0.$$

$$q_j(x) \sum_i c_i q_i(x) = q_j(x) \cdot 0$$

$$\sum_i c_i q_i(x) q_j(x) = 0$$

$$\int_D \sum_i c_i q_i(x) q_j(x) \, dx = \int_D 0 \, dx$$

$$c_j \lambda_j = 0$$

$$\lambda_j = 0$$

However, this goes against the rule that for orthogonal functions, $\langle q_i, q_j \rangle = \lambda_i > 0$ when $i = j$. Therefore, in this case $\{q_i(x)\}$ can't be a set of orthogonal functions. Having completed a proof by contrapositive, it's possible to infer that the original statement that if $\{q_i(x)\}$ is orthogonal then it must also be linearly independent is true. ■

(b) Prove that the integrated mean squared error is the sum of the integrated variance and the integrated squared bias, that is,

$$IMSE(\hat{f}) = IV(\hat{f}) + ISB(\hat{f}).$$

$$IMSE(\hat{f}) = \int_D E\left[\left(\hat{f}(x) - f(x)\right)^2\right] dx$$

$$= \int_D E\left[\left(\hat{f}(x) - E\left(\hat{f}(x)\right) + E\left(\hat{f}(x)\right) - f(x)\right)^2\right] dx$$

$$= \int_D E\left\{\left[\hat{f}(x) - E\left(\hat{f}(x)\right)\right]^2 + 2\left[\hat{f}(x) - E\left(\hat{f}(x)\right)\right]\left[E\left(\hat{f}(x)\right) - f(x)\right] + \left[E\left(\hat{f}(x)\right) - f(x)\right]^2\right\} dx$$

$$= \underbrace{\int_D E\left\{\left[\hat{f}(x) - E(\hat{f}(x))\right]^2\right\} dx}_{IV(\hat{f})} + 2\int_D E\{[\hat{f}(x) - E(\hat{f}(x))][E(\hat{f}(x)) - f(x)]\} dx + \underbrace{\int_D E\left\{\left[E(\hat{f}(x)) - f(x)\right]^2\right\} dx}_{ISB(\hat{f})}$$

The part in the center within the integral can be written as follows:

$$[E(\hat{f}(x)) - E(\hat{f}(x))][E(\hat{f}(x)) - f(x)] = 0 \cdot \left[E\left(\hat{f}(x)\right) - f(x)\right] = 0$$

$$\therefore IMSE(\hat{f}) = IV(\hat{f}) + ISB(\hat{f}) \blacksquare$$

3. Let $\{q_k : k = 1, \cdots, m\}$ be a set of orthogonal functions. Show that

$$\left\|\sum_{k=1}^{m} q_k\right\|^2 = \sum_{k=1}^{m} \|q_k\|^2.$$

$$\left\|\sum_{k=1}^{m} q_k\right\|^2 = \int_D \left(\sum_{k=1}^{m} q_k(x)\right)^2 dx$$

$$= \int_D \sum_{i=1}^{m} q_i^2(x) + \sum_{\substack{i,j=1 \\ i \neq j}}^{m} q_i(x) q_j(x)\, dx$$

$$= \sum_{i=1}^{m} \int_D q_i^2(x)\, dx + \sum_{\substack{i,j=1 \\ i \neq j}}^{m} \int_D q_i(x) q_j(x)\, dx$$

$$= \sum_{i=1}^{m} \|q_i(x)\|^2 + 0 = \sum_{k=1}^{m} \|q_k\|^2 \ \blacksquare$$

What is the common value of the expression above if the q_k are orthonormal? In these expressions, $\|\cdot\|$ represents an L_2 norm. Would a similar equation hold for a general L_p norm?

Given that q_k are orthonormal, then $\|q_k\|^2 = 1 \; \forall \; k$. Therefore, the value would be $\sum_{k=1}^{m} \|q_k\|^2 = \boxed{m}$. If there was an alternate L_p norm such as L_3, there's no guarantee that the above situation would hold. The reason is that the inner product cancelling out only works for the product of two different q_i and q_j functions. If there's the product of $q_i q_j q_k$ for example, the solution is not certain since there's no rule saying whether or not they cancel out properly the way that two orthogonal functions do.

4. (a) Use the Gram-Schmidt orthogonalization process as described in Lecture 10C to derive the first four orthonormal Chebychev polynomials. Note that the range is $[-1, 1]$ and the weight function is $w(x) = (1 - x^2)^{1/2}$. (Please note that the set is to be orthonormal and not just orthogonal.)

Let $\tilde{q}_i$, $i = 0, \cdots, 3$ denote the orthogonalized set of functions. Also, let the linearly independent set $\{q_i(x)\}$, $i = 0, \cdots, 3$ be $\{1, x, x^2, x^3\}$ and

$$\tilde{q}_k = q_k - \sum_{i=1}^{k-1} \frac{\langle \tilde{q}_i, q_k \rangle}{\langle \tilde{q}_i, \tilde{q}_i \rangle} \cdot \tilde{q}_i.$$

Then from the Gram-Schmidt orthogonalization process these are the first four polynomials.

$$\tilde{q}_0 = q_0 = 1$$

$$\tilde{q}_1 = x - \frac{\langle 1, x \rangle}{\langle 1,1 \rangle} \cdot 1 = x$$

$$\tilde{q}_2 = x^2 - \frac{\langle 1, x^2 \rangle}{\langle 1,1 \rangle} \cdot 1 - \frac{\langle x, x^2 \rangle}{\langle x, x \rangle} \cdot x = x^2 - \frac{1}{4}$$

$$\tilde{q}_3 = x^3 - \frac{\langle 1, x^3 \rangle}{\langle 1,1 \rangle} \cdot 1 - \frac{\langle x, x^3 \rangle}{\langle x, x \rangle} \cdot x - \frac{\left\langle \left(x^2 - \frac{1}{4}\right), x^3 \right\rangle}{\left\langle \left(x^2 - \frac{1}{4}\right), \left(x^2 - \frac{1}{4}\right) \right\rangle} \cdot \left(x^2 - \frac{1}{4}\right) = x^3 - \frac{x}{2}$$

The above derivations were calculated from the following:

$$\langle q_i, q_j \rangle = \int_{-1}^{1} q_i(x) q_j(x) w(x) dx = \int_{-1}^{1} q_i(x) q_j(x) \sqrt{1 - x^2} dx$$

$$\langle 1,1 \rangle = \int_{-1}^{1} \sqrt{1 - x^2} dx = \cdots = \frac{\pi}{2}$$

$$\langle 1, x \rangle = \int_{-1}^{1} x \sqrt{1 - x^2} dx = \cdots = 0$$

$$\langle x, x \rangle = \langle 1, x^2 \rangle = \int_{-1}^{1} x^2 \sqrt{1 - x^2} dx = \cdots = \frac{\pi}{8}$$

$$\langle x, x^2 \rangle = \langle 1, x^3 \rangle = \int_{-1}^{1} x^3 \sqrt{1-x^2}\,dx = \cdots = 0$$

$$\langle x, x^3 \rangle = \int_{-1}^{1} x^4 \sqrt{1-x^2}\,dx = \cdots = \frac{\pi}{16}$$

$$\langle \left(x^2 - \frac{1}{4}\right), x^3 \rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{4}\right) x^3 \sqrt{1-x^2}\,dx = \cdots = 0$$

$$\langle \left(x^2 - \frac{1}{4}\right), \left(x^2 - \frac{1}{4}\right) \rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{4}\right)^2 \sqrt{1-x^2}\,dx = \cdots = \frac{\pi}{32}$$

The above inner products were calculated using an online calculator: www.integral-calculator.com Permission was given by the professor to use this tool as long as credit is given. The reason is that the above calculations require the use of trigonometric identities and so the entire process would be highly time consuming to complete by hand.

Now that the orthogonal set $\{\tilde{q}_i(x)\}$, $i = 0, \cdots, 3$ have been calculated, they need to be normalized to create the orthonormal set. This calculation is performed by dividing by the magnitude, $\left\{\frac{\tilde{q}_i(x)}{\|\tilde{q}_i(x)\|}\right\}$. Denote the orthonormal set to be $\{\hat{q}_i(x)\}$, $i = 0, \cdots, 3$.

$$\hat{q}_0 = \frac{\tilde{q}_0}{\sqrt{\langle \tilde{q}_0, \tilde{q}_0 \rangle}} = \frac{1}{\sqrt{\int_{-1}^{1} \sqrt{1-x^2}\,dx}} = \boxed{\left(\frac{\pi}{2}\right)^{-1/2}}$$

$$\hat{q}_1 = \frac{\tilde{q}_1}{\sqrt{\langle \tilde{q}_1, \tilde{q}_1 \rangle}} = \frac{x}{\sqrt{\int_{-1}^{1} x^2\sqrt{1-x^2}\,dx}} = \frac{x}{\sqrt{\frac{\pi}{8}}} = \boxed{x\left(\frac{\pi}{8}\right)^{-1/2}}$$

$$\hat{q}_2 = \frac{\tilde{q}_2}{\sqrt{\langle \tilde{q}_2, \tilde{q}_2 \rangle}} = \frac{\left(x^2 - \frac{1}{4}\right)}{\sqrt{\int_{-1}^{1} \left(x^2 - \frac{1}{4}\right)^2 \sqrt{1-x^2}\,dx}} = \frac{\left(x^2 - \frac{1}{4}\right)}{\sqrt{\frac{\pi}{32}}} = \boxed{\left(x^2 - \frac{1}{4}\right)\left(\frac{\pi}{32}\right)^{-1/2}}$$

$$\hat{q}_3 = \frac{\tilde{q}_3}{\sqrt{\langle \tilde{q}_3, \tilde{q}_3 \rangle}} = \frac{\left(x^3 - \frac{x}{2}\right)}{\sqrt{\int_{-1}^{1} \left(x^3 - \frac{x}{2}\right)^2 \sqrt{1-x^2}\,dx}} = \frac{\left(x^3 - \frac{x}{2}\right)}{\sqrt{\frac{\pi}{128}}} = \boxed{\left(x^3 - \frac{x}{2}\right)\left(\frac{\pi}{128}\right)^{-1/2}}$$

NOTE: The same online integral calculator was used for the above equations.

(b) Posted on the course Blackboard is the data set Orthogonal.txt containing 1000 observations from a $N(0, .3^2)$ distribution (standard deviation of $.3$). Use the polynomials derived in part (a) to estimate the density f. Do you think that this is a good estimate of f? Be sure to explain why or why not.

In this problem, to approximate f it was necessary to find the values of c_k in

$$f(x) = \sum_{k=0}^{3} c_k q_k(x).$$

These values were approximated using,

$$\hat{c}_k = \frac{1}{n}\sum_{i=1}^{n} q_k(x_i) g(x_i).$$

Here, $g(x) = 1$ since the data comes from a normal density without any other function attached to it. Using R, below are the estimated values of $\hat{c}_k$:

$\hat{c}_0$	$\hat{c}_1$	$\hat{c}_2$	$\hat{c}_3$
0.7978846	-0.003898395	-0.5041277	0.008430499

Below in Figure 1 is a plot of the estimated function in black and the true function in red. The true function was plot by using the `dnorm` function in R with the corresponding parameters of 0 and 0.3 for the mean and standard deviation. It makes sense to think that it wouldn't estimate f perfectly since there are only 4 orthonormal polynomials compared to the theoretical limit of infinity.

Looking at the graph, it seems that the estimated function does have part of the characteristic bell curve appearance, without rounding off to the sides. Instead, the values continue towards around -15. This is understandable since this is an estimated function based off some sample data. Therefore, it doesn't necessarily follow that it would perfectly fit the shape of a probability density function that has a bell shape going from $-\infty$ to ∞ which are based off expressing extreme outliers that are unseen in the data.

Also, the estimated function does happen to cover most of the area under the true function. However, it's slightly shorter in the center and higher on the sides, giving it properties that would make it seem that it has something like having heavier tails. Overall, it's a decent estimate of the true function f, however, with more orthonormal polynomials and a larger dataset it would make sense to think that the estimate would improve.

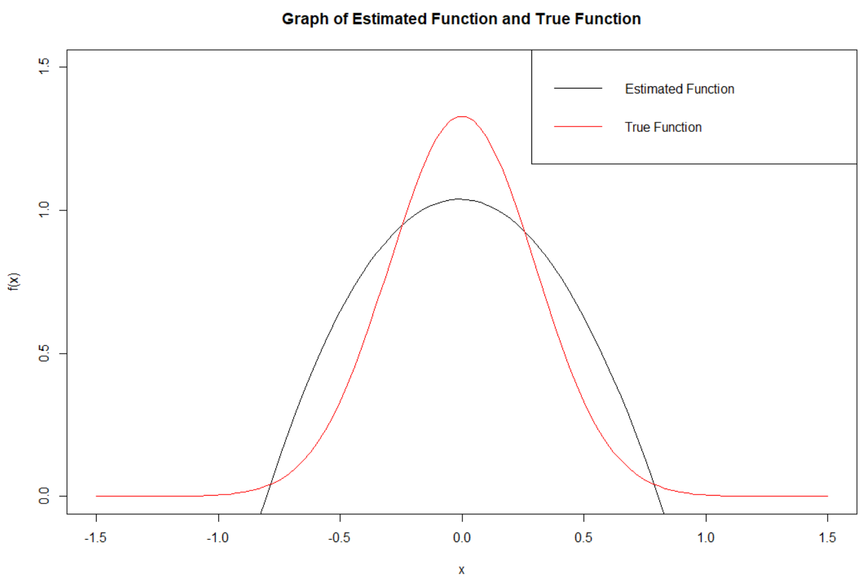

Figure 1

5. (a) Which properties of a natural cubic spline does the following function possess and which does it not possess?

$$f(x) = \begin{cases} (x+1) + (x+1)^3 & x \in [-1,0] \\ 4 + (x-1) + (x-1)^3 & x \in (0,1] \end{cases}$$

A natural cubic spline has the following properties:

1. $f_{i-1}(t_i) = y_i = s_i(t_i)$
2. $s'_{i-1}(t_i) = s'_i(t_i)$
3. $s''_{i-1}(t_i) = s''_i(t_i) = z_i$

Additionally, the cubic spline is a polynomial of degree ≤ 3 and is continuous.

The above function $f(x)$ is a polynomial of less than or equal to 3.

Also, the only break point in the function is at 0, and the limits of the function as x approaches from left and right are both 2. Therefore, it's a continuous function. So, the first property holds.

Let $s_0(x) = (x+1) + (x+1)^3$ and $s_1(x) = 4 + (x-1) + (x-1)^3$. Then the following are the set of first and second derivatives for $s_0(x)$ and $s_1(x)$,

$$\begin{matrix} s'_0\,(x) = 1 + 3(x+1)^2 & s'_1(x) = 1 + 3(x-1)^2 \\ s''_0(x) = 6(x+1) & s''_1(x) = 6(x-1) \end{matrix}.$$

From here, it's clear to see that $s'_{i-1}(t_i) = s'_i(t_i)$ and $s''_{i-1}(t_i) = s''_i(t_i) = z_i$ don't hold, because $s'_0\,(x) \neq s'_1(x)$ and $s''_0\,(x) \neq s''_1(x)$. So here the second and third property don't hold.

It was mentioned too that the function is a polynomial with degree ≤ 3. To be specific, $s_0(x)$ and $s_1(x)$ are both polynomials of degree 3. Therefore, $s_0''(x)$ and $s_1''(x)$ are linear functions (another property of cubic splines). However, z_i doesn't exist since the second derivatives are different. Then it follows that $s_i''(t_i) = z_i$ and $s_{i+1}''(t_i) = z_{i+1}$ is irrelevant.

(b) Find a natural cubic spline function whose knots are $1, 2, 3,$ and 4 that takes values $f(1) = 1$, $f(2) = 1/2$, $f(3) = 1/3$, and $f(4) = 1/4$.

$$\begin{bmatrix} u_1 & h_1 \\ h_1 & u_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$v_1 = u_1 z_1 + h_1 z_2$$

$$v_1 = b_1 - b_0 = 2$$

$$b_0 = \frac{6}{h_i}(y_{i+1} - y_i) = 6\left(-\frac{1}{2}\right) = -3$$

$$b_1 = \cdots = -1$$

$$u_1 = 2(1+1) = 4$$

$$\rightarrow 2 = 4z_1 + z_2$$

$$v_2 = h_1 z_1 + u_2 z_2$$

$$b_2 = 6\left(\frac{1}{4} - \frac{1}{3}\right) = -\frac{1}{2}$$

$$(b_2 - b_1) = z_1 + 4z_2$$

$$u_2 = 2(1+1) = 4$$

$$\rightarrow \frac{1}{2} = z_1 + 4z_2$$

$$z_1 = \frac{1}{2} - 4z_2$$

$$4\left(\frac{1}{2} - 4z_2\right) + z_2 = 2$$

$$2 - 16z_2 + z_2 = 2$$

$$z_2 = 0, z_1 = \frac{1}{2}; z_0 = z_3 = 0$$

$$s_0(x) = 0 + \frac{\frac{1}{2}}{6}(x-1)^3 + \left(\frac{1}{2} - \frac{\frac{1}{2}}{6}\right)(x-1) + (1-0)(2-x)$$

$$= \frac{1}{2}(x-1)^3 + \frac{5}{12}(x-1) + (2-x)$$

$$s_1(x) = \frac{\frac{1}{2}}{6}(3-x)^3 + 0 + \left(\frac{1}{3} - 0\right)(x-2) + \left(\frac{1}{2} - \frac{1}{12}\right)(3-x)$$

$$= \frac{1}{12}(3-x)^3 + \frac{1}{3}(x-2) + \frac{5}{12}(3-x)$$

$$s_2(x) = 0 + 0 + \left(\frac{1}{4} - 0\right)(x-3) + \frac{1}{3}(4-x) = \frac{1}{4}(x-3) + \frac{1}{3}(4-x)$$

Then the natural cubic spline is as follows:

$$S(x) = \begin{cases} \frac{1}{2}(x-1)^3 + \frac{5}{12}(x-1) + (2-x) & x \in [1,2] \\ \frac{1}{12}(3-x)^3 + \frac{1}{3}(x-2) + \frac{5}{12}(3-x) & x \in [2,3] \\ \frac{1}{4}(x-3) + \frac{1}{3}(4-x) & x \in [3,4] \end{cases}$$

6. (Extra Credit) Prove that the Fourier coefficients form the finite expansion in basis functions with the minimum mean squared error, that is prove

$$\left\| f - \sum_{k=0}^{j} c_k q_k \right\|^2 \le \left\| f - \sum_{k=0}^{j} a_k q_k \right\|^2$$

where $\{c_k = \langle f, q_k \rangle\}$ are the Fourier coefficients and $\{a_k\}$ are any other constants. Assume that the basis set is orthonormal and note that $\|\cdot\|$ is the L_2 norm. Also, you may find it useful to use the results of Problem 3. (Hint: Write $\|f - a_0 q_0\|^2$ a function of a_0, $\langle f, f \rangle - 2a_0\langle f, q_0 \rangle + a_0^2\langle q_0, q_0 \rangle$, differentiate, set to zero for the minimum, and determine $a_0 = c_0$. This same approach can be done in multidimensions for $a_0, a_1, a_2, \cdots, a_k$ or else induction can be used from a_1 on.)

$$g(\vec{a}) = \left\| f - \sum_{k=0}^{j} a_k q_k \right\|^2$$

$$= \int_D \left(f(x) - \sum_{k=0}^{j} a_k q_k(x) \right)^2 dx$$

$$= \int_D f(x)^2 - 2f(x)\sum_{k=0}^{j} a_k q_k(x) + \left(\sum_{k=0}^{j} a_k q_k(x)\right)^2 dx$$

$$= \|f\|^2 - 2\int_D f(x)\sum_{k=0}^{j} a_k q_k(x)\,dx + \int_D \sum_{k=0}^{j}\big(a_k q_k(x)\big)^2 + \sum_{\substack{i,k=0\\ i\neq k}}^{j} a_i a_k q_i(x) q_k(x)\,dx$$

$$= \|f\|^2 - 2\sum_{k=0}^{j} a_k \int_D f(x) q_k(x)\,dx + \sum_{k=0}^{j} a_k^2$$

$$= \|f\|^2 - 2\sum_{k=0}^{j} a_k\, c_k + \sum_{k=0}^{j} a_k^2$$

$$\frac{\partial g(\vec{a})}{\partial a_k} = \frac{\partial}{\partial a_k}\left[\|f\|^2 - 2\sum_{k=0}^{j} a_k\, c_k + \sum_{k=0}^{j} a_k^2\right]$$

$$= 0 - 2c_k + 2a_k = 2a_k - 2c_k$$

$$\frac{\partial g(\vec{a})}{\partial a_k} \overset{set\ to}{\triangleq} 0$$

$$\rightarrow 2a_k = 2c_k$$

$$a_k = c_k \;\forall\, k$$

Therefore, the value of a_k for $k = 0, \cdots, j$ that minimizes $\left\|f - \sum_{k=0}^{j} a_k q_k\right\|^2$ is c_k. This proves the statement that $\left\|f - \sum_{k=0}^{j} c_k q_k\right\|^2 \leq \left\|f - \sum_{k=0}^{j} a_k q_k\right\|^2$. ■

Appendix

```
### Q4 part(b)
orthogonal <- read.csv('Orthogonal.txt', header = FALSE)
orthogonal <- orthogonal[[1]]
hist(orthogonal)
plot(1:length(orthogonal), orthogonal)

q0 <- function(x) {
  (pi / 2)^(-1/2)
}

q1 <- function(x) {
  x * ((pi / 8)^(-1 / 2))
```

```r
}

q2 <- function(x) {
  ((x^2) - (1 / 4)) / (sqrt(pi / 32))
}

q3 <- function(x) {
  ((x^3) - (x / 2)) / sqrt(pi / 128)
}

c0 <- mean(q0(orthogonal))
c1 <- mean(q1(orthogonal))
c2 <- mean(q2(orthogonal))
c3 <- mean(q3(orthogonal))

f <- function(x) {
  # c0 * q0(x) + c1 * q1(x) + c2 * q2(x)
  c0 * q0(x) + c1 * q1(x) + c2 * q2(x) + c3 * q3(x)
}

xs <- seq(-1.5, 1.5, length.out = 1e2)
plot(xs, f(xs), type = 'l', ylim = c(0, 1.5),
     main = 'Graph of Estimated Function and True Function',
     xlab = 'x', ylab = 'f(x)')
lines(xs, dnorm(xs, mean = 0, sd = 0.3), col = 'red')
legend("topright", legend = c('Estimated Function', 'True Function'),
       col = c('black', 'red'), lty = c(1,1))

### Q5
# part (a)
f <- function(x) {
  if ((x >= -1) & (x <= 0)) {
    (x + 1) + (x + 1)^3
  } else if ((x > 0) & (x <= 1)) {
    4 + (x - 1) + (x - 1)^3
  } else {
    0
  }
}

xs <- seq(-1.1, 1.1, length.out = 1e4)
f_vec <- Vectorize(f)
plot(xs, f_vec(xs), type = 'l')

# part (b)
s0 <- function(x) {
  (1 / 12) * ((x - 1)^3) + (5 / 12) * (x - 1) + (2 - x)
}

s1 <- function(x) {
  (1 /12) * ((3 - x)^3) + (1 / 3) * (x - 2) + (5 / 12) * (3 - x)
}

s2 <- function(x) {
  (1 / 4) * (x - 3) + (1 / 3) * (4 - x)
```

```r
}

s0(1)
s0(2) == s1(2)
s1(3) == s2(3)
s2(4)

b = 0.01
x = c(seq(1,2,b), seq(2,3,b), seq(3,4,b))
y = c(s0(seq(1,2,b)), s1(seq(2,3,b)), s2(seq(3,4,b)))
plot(x,y, type = 'l')
abline(v = 1, col='red')
abline(v = 2, col='red')
abline(v = 3, col='red')
abline(v = 4, col='red')
```