Numerical Analysis II: Homework 2
Spring 2018
Due Feb 9
Lindsay Eddy

This homework assignment covers the materials on approximation theory. For the computer problems, please attach your computer codes to the homework as well.

1. [10 points] Problem 4 at the end of Chapter 2 of the textbook [1, p. 119].

   ...Find the optimal $c = c_p$ for $p = \infty$, $p = 2$, and $p = 1$ and determine $E_p(c_p)$ for each of these p-values.

   - $\underline{p = \infty}$:

     $E_\infty(c) = \|t^\alpha - c\|_\infty$, i.e., the least upper bound of the function $t^\alpha - c$ on the interval $[0, 1]$. In other words, the c that minimizes $E_\infty(c)$ minimizes $\max_{t\in[0,1]}(t^\alpha - c)$. This value is minimized when
     $$c = \frac{\max(t^\alpha) + \min(t^\alpha)}{2}$$
     where $t \in [0, 1]$. Since $\alpha > 0$, t^α is monotonically increasing on $[0, 1]$, so $\min(t^\alpha) = 0^\alpha = 0$ and $\max(t^\alpha) = 1^\alpha = 1$.

     Therefore the c that minimizes $E_\infty(c)$ is $\frac{1}{2}$, and $E_\infty(1/2) = \frac{1}{2}$.

   - $\underline{p = 2}$:

     $$E_2(c) = \|t^\alpha - c\|_2 = \left(\int_0^1 (t^\alpha - c)^2 \mathrm{d}t\right)^{1/2}.$$

     Since the function $f(x) = x^{1/2}$ is monotonically increasing on the interval $[0, 1]$, $E_2(c)$ is minimized when
     $$\int_0^1 (t^\alpha - c)^2 \mathrm{d}t \tag{1}$$
     is minimized. Thus the $c \in [0, 1]$ that minimizes (1) also minimizes $E_2(c)$ on $[0, 1]$.

     To minimize (1):
     $$\begin{aligned}\int_0^1 (t^\alpha - c)^2 \mathrm{d}t &= \int_0^1 (t^\alpha - 2ct^\alpha + c^2)\mathrm{d}t \\ &= \left(\frac{t^{2\alpha+1}}{2\alpha+1} - 2c\frac{t^{\alpha+1}}{\alpha+1} + tc^2\right)\Bigg|_{t=0}^{t=1} \\ &= \frac{1}{2\alpha+1} - \frac{2c}{\alpha+1} + c^2\end{aligned}$$

     So
     $$\frac{\mathrm{d}}{\mathrm{d}c}\left(\int_0^1 (t^\alpha - c)^2 \mathrm{d}t\right) = \frac{\mathrm{d}}{\mathrm{d}c}\left(\frac{1}{2\alpha+1} - \frac{2c}{\alpha+1} + c^2\right) = -\frac{2}{\alpha+1} + 2c$$

     Setting this equal to zero we have $c = 1/(\alpha + 1)$.

     Also,
     $$\frac{\mathrm{d}^2}{\mathrm{d}c^2}\left(\int_0^1 (t^\alpha - c)^2 \mathrm{d}t\right) = \frac{\mathrm{d}}{\mathrm{d}c}\left(-\frac{2}{\alpha+1} + 2c\right) = 2.$$

     This is always positive, so $c = 1/(\alpha + 1)$ in fact minimizes (1).

$$
\begin{aligned}
E_2(1/(\alpha+1)) &= \left\| t^\alpha - \frac{1}{\alpha+1} \right\|_2 \\
&= \left(\int_0^1 \left(t^\alpha - \frac{1}{\alpha+1} \right)^2 \mathrm{d}t \right)^{1/2} \\
&= \left(\frac{1}{2\alpha+1} - \frac{2}{(\alpha+1)^2} + \frac{1}{(\alpha+1)^2} \right)^{1/2} \\
&= \left(\frac{\alpha^2}{(2\alpha+1)(\alpha+1)^2} \right)^{1/2} \\
&= \frac{\alpha}{(2\alpha+1)^{1/2}(\alpha+1)}
\end{aligned}
$$

- $\underline{p = 1}$:

$$
E_1(c) = \|t^\alpha - c\|_1 = \int_0^1 |t^\alpha - c| \mathrm{d}t
$$

So

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}c} E_1(c) &= \frac{\mathrm{d}}{\mathrm{d}c} \left(\int_0^1 |t^\alpha - c| \mathrm{d}t \right) \\
&= \frac{\mathrm{d}}{\mathrm{d}c} \left(\int_0^{c^{1/\alpha}} -(t^\alpha - c) \mathrm{d}t \ + \ \int_{c^{1/\alpha}}^1 (t^\alpha - c) \mathrm{d}t \right) \\
&= \int_0^{c^{1/\alpha}} \frac{\mathrm{d}}{\mathrm{d}c}(c - t^\alpha) \mathrm{d}t + \int_{c^{1/\alpha}}^1 \frac{\mathrm{d}}{\mathrm{d}c}(t^\alpha - c) \mathrm{d}t \\
&= \int_0^{c^{1/\alpha}} \mathrm{d}t + \int_{c^{1/\alpha}}^1 -\mathrm{d}t \\
&= 2c^{1/\alpha} - 1
\end{aligned}
$$

Setting this equal to zero we have $c = 2^{-\alpha}$.

Also,

$$
\frac{\mathrm{d}^2}{\mathrm{d}c^2} E_1(c) = \frac{\mathrm{d}}{\mathrm{d}c}(2c^{1/\alpha} - 1) = \frac{2\,c^{(1/\alpha)-1}}{\alpha}
$$

Since $\alpha > 0$, $\frac{\mathrm{d}^2}{\mathrm{d}c^2} E_1(c)$ is positive for all $c \in [0,1]$. Thus $c = 2^{-\alpha}$ in fact minimizes $E_1(c)$ on the interval $[0,1]$.

$$
\begin{aligned}
E_1(2^{-\alpha}) &= \int_0^1 |t^\alpha - 2^{-\alpha}| \mathrm{d}t \\
&= 2(2^{-\alpha(\frac{1}{\alpha}+1)}) - \frac{2(2^{-\alpha(\frac{1}{\alpha}+1)})}{\alpha+1} + \frac{1}{\alpha+1} - 2^{-\alpha} \\
&= \frac{1 - 2^{-\alpha}}{\alpha+1}
\end{aligned}
$$

2. [5 points] Consider $L^2_w[a,b]$ where $w \geq 0$ is a weight function on $[a,b]$. This space is endowed with inner product $(f,g) = \int_a^b f(x)g(x)w(x)\,dx$, and norm $\|f\|_{2,w} = (f,f)^{1/2}$. Suppose $\{\pi_j\}_{i=1}^n$ is a linearly *dependent* set in $L^2_w[a,b]$. Prove that the matrix $\mathbf{A}$ with entries $A_{ij} = (\pi_i, \pi_j)$ is singular.

Since $\{\pi_j\}_{i=1}^n$ is a linearly dependent set in $L^2_w[a,b]$, there exists an n x 1 vector $\boldsymbol{x}$ such that $\sum_{j=1}^n x_j \pi_j = 0$ where some $x_k \neq 0$ $(k \in 1, \ldots, n)$.
Thus there exists an n x 1 vector $\boldsymbol{x}$ such that

$$(Ax)_{ij} = \sum_{j=1}^n x_j(\pi_i, \pi_j) = (\pi_i, \textstyle\sum_{j=1}^n x_j\pi_j) = (\pi_i, 0) = 0\,(\pi_i, 0) = 0$$

for all i, j, where some $x_k \neq 0$. $x_k \neq 0 \Rightarrow \boldsymbol{x} \neq \mathbf{0}$ and $(Ax)_{ij} = 0 \;\forall\; i,j \Rightarrow \mathbf{A}\boldsymbol{x} = \mathbf{0}$. Therefore, there exists an $\boldsymbol{x} \neq \mathbf{0}$ such that $\mathbf{A}\boldsymbol{x} = \mathbf{0}$.
$\therefore$ $\mathbf{A}$ is singular.

3. [10 points] Let $f \in L^2_w[a,b]$ and let Φ_n be a finite-dimensional subspace of $L^2_w[a,b]$. Consider the least squares approximation problem of finding φ_n^* such that $\|f - \varphi_n^*\|_{2,w} = \min_{\varphi\in\Phi_n} \|f - \varphi\|_{2,w}$. Putting a basis on Φ_n we can derive the corresponding normal equations and compute the solution to the least squares approximation problem. In this exercise we prove the uniqueness of the least-squares approximation, independently of the choice of basis, using an elementary argument.

   (a) Derive the identity[1]

$$\|f+g\|^2_{2,w} + \|f-g\|^2_{2,w} = 2\|f\|^2_{2,w} + 2\|g\|^2_{2,w}, \quad \text{for all } f, g \in L^2_w[a,b],$$

   using properties of the inner product and the corresponding induced norm.

(Let $\|\cdot\|$ denote $\|\cdot\|_{2,w}$ and let $(\cdot,\cdot)$ denote $(\cdot,\cdot)_{2,w}$.)

$$\begin{aligned}\|f+g\|^2 + \|f-g\|^2 &= (f+g, f+g) + (f-g, f-g)\\ &= \|f\|^2 + 2(f,g) + \|g\|^2 + \|f\|^2 - 2(f,g) + \|g\|^2\\ &= 2\|f\|^2 + 2\|g\|^2\end{aligned}$$

   (b) Show that if $f \neq g$ and $\|f\|_{2,w} = \|g\|_{2,w} = 1$, then $\|\frac{1}{2}(f+g)\|_{2,w} < 1$.

(Let $\|\cdot\|$ denote $\|\cdot\|_{2,w}$ and let $(\cdot,\cdot)$ denote $(\cdot,\cdot)_{2,w}$.)

$$\begin{aligned}\|\tfrac{1}{2}(f+g)\|^2 &= (\tfrac{1}{2}(f+g), \tfrac{1}{2}(f+g))\\ &= \tfrac{1}{4}\|f+g\|^2\\ &= \tfrac{1}{2}\|f\|^2 + \tfrac{1}{2}\|g\|^2 - \tfrac{1}{4}\|f-g\|^2\\ &= 1 - \tfrac{1}{4}\|f-g\|^2\end{aligned}$$

Since $f \neq g$, $\|f-g\| > 0$ (by the positivity and non-degeneracy properties of norms). Thus $\|f-g\|^2 > 0$, and consequently $\|\frac{1}{2}(f+g)\|^2 < 1$, which implies $\|\frac{1}{2}(f+g)\| < 1$.

   (c) Show that solution $\varphi^* \in \Phi_n$ of the least squares problem is unique. Prove this by assuming there are two functions φ^* and ψ^* in Φ_n, with $\varphi^* \neq \psi^*$, such that $\|f - \varphi^*\|_{2,w} = \|f - \psi^*\|_{2,w} = \min_{\varphi\in\Phi_n} \|f-\varphi\|_{2,w}$ and arrive at a contradiction. [Hint: consider $\|f - \frac{1}{2}(\varphi^* + \psi^*)\|_{2,w}$.]

[1] This is known as the parallelogram identity; it holds for any norm induced by an inner product.

(Let $\|\cdot\|$ denote $\|\cdot\|_{2,w}$ and let $(\cdot,\cdot)$ denote $(\cdot,\cdot)_{2,w}$.)

Let φ^*, $\psi^* \in \Phi_n$ such that $\|f-\varphi^*\| = \|f-\psi^*\| = \min_{\varphi\in\Phi_n}\|f-\varphi\|$. Assume $\varphi^* \neq \psi^*$.

$$\begin{aligned}2\|f-\tfrac{1}{2}(\varphi^*+\psi^*)\|^2 &= 2\|f\|^2 - 2(f,\varphi^*) - 2(f,\psi^*) + \tfrac{1}{2}\|\varphi^*+\psi^*\|^2\\ &= 2\|f\|^2 - 2(f,\varphi^*) - 2(f,\psi^*) + \tfrac{1}{2}(2\|\varphi^*\|^2 + 2\|\psi^*\|^2 - \|\varphi^*-\psi^*\|^2)\\ &= 2\|f\|^2 - 2(f,\varphi^*) - 2(f,\psi^*) + \|\varphi^*\|^2 + \|\psi^*\|^2 - \tfrac{1}{2}\|\varphi^*-\psi^*\|^2\end{aligned}$$

Since $\varphi^* \neq \psi^*$, $\|\varphi^*-\psi^*\| > 0$ (by the positivity and non-degeneracy properties of norms). Therefore,

$$\begin{aligned}2\|f-\tfrac{1}{2}(\varphi^*+\psi^*)\|^2 &= 2\|f\|^2 - 2(f,\varphi^*) - 2(f,\psi^*) + \|\varphi^*\|^2 + \|\psi^*\|^2 - \tfrac{1}{2}\|\varphi^*-\psi^*\|^2\\ &< 2\|f\|^2 - 2(f,\varphi^*) - 2(f,\psi^*) + \|\varphi^*\|^2 + \|\psi^*\|^2\\ &= \|f\|^2 - 2(f,\varphi^*) + \|\varphi^*\|^2 + \|f\|^2 - 2(f,\psi^*) + \|\psi^*\|^2\\ &= \|f-\varphi^*\|^2 + \|f-\psi^*\|^2\\ &= 2\|f-\varphi^*\|^2 \qquad \text{(Since } \|f-\varphi^*\| = \|f-\psi^*\|\text{)}.\end{aligned}$$

We have shown

$$2\|f-\tfrac{1}{2}(\varphi^*+\psi^*)\|^2 < 2\|f-\varphi^*\|^2$$

Since all norms are non-negative,

$$\|f-\tfrac{1}{2}(\varphi^*+\psi^*)\| < \|f-\varphi^*\| = \min_{\varphi\in\Phi_n}\|f-\varphi\|.$$

Since $\frac{1}{2}(\varphi^*+\psi^*) \in \Phi_n$, this is a contradiction. Thus $\varphi^* = \psi^*$. That is, the solution $\varphi^* \in \Phi_n$ is unique.

4. [20 points] Consider the Legendre polynomials given by, $\pi_n(x) = \dfrac{1}{2^n n!}\dfrac{d^n}{dx^n}\left[(x^2-1)^n\right]$, for $n \geq 0$. These polynomials form a complete orthogonal system in $L^2[-1,1]$, [2] where $L^2[-1,1]$ is equipped with the inner product $(f,g) = \int_{-1}^{1} f(x)g(x)\,dx$, and norm $\|\cdot\|_2 = (\cdot,\cdot)^{1/2}$. Specifically, we have

$$(\pi_n,\pi_m) = \frac{2}{2n+1}\delta_{nm}. \tag{2}$$

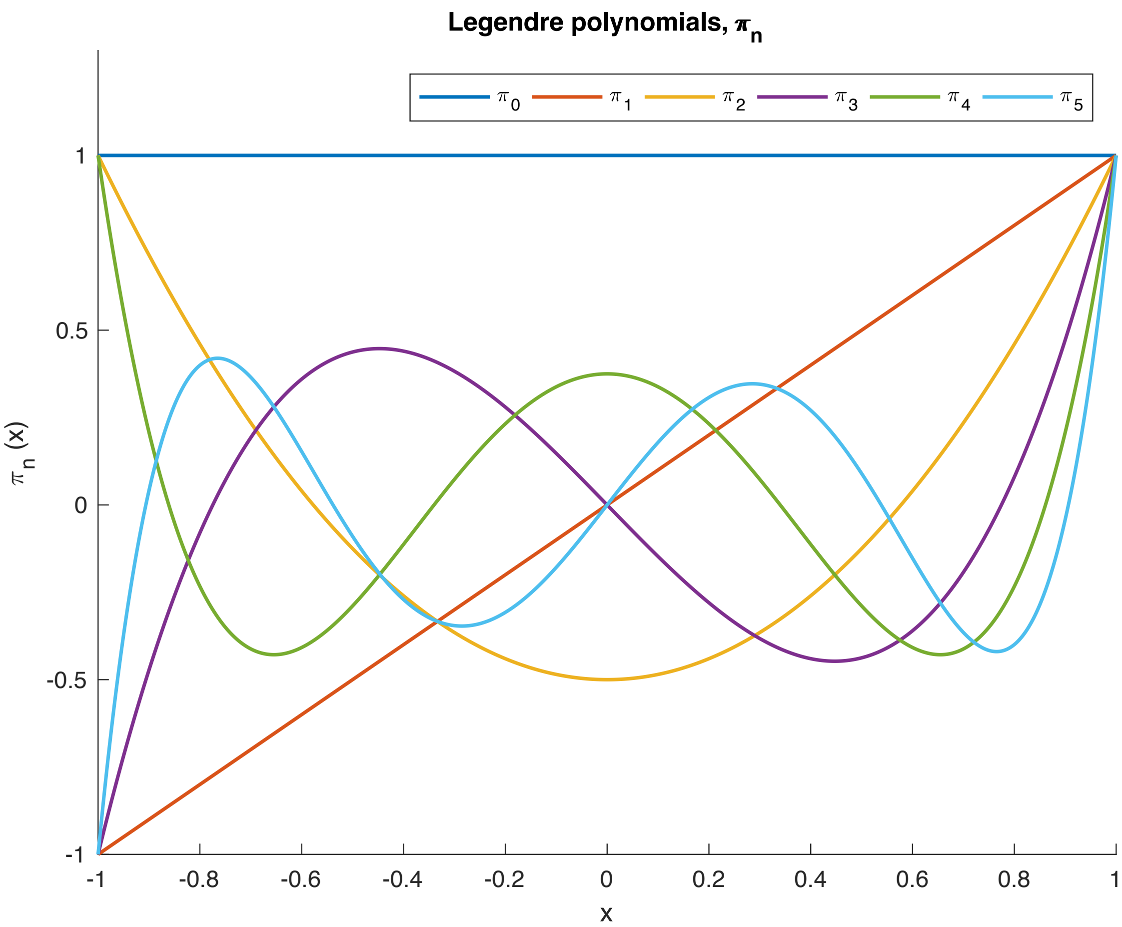
(a) The Legendre polynomials satisfy, $\pi_0 \equiv 1$, $\pi_1(x) = x$, and

$$\pi_{n+1}(x) = \frac{2n+1}{n+1}x\pi_n(x) - \frac{n}{n+1}\pi_{n-1}(x), \quad n \geq 1.$$

This relation provides an easy way of generating the Legendre polynomials. This is implemented in the computer code `legpoly`, which you can download from the Moodle page for the course. Use this Matlab function to compute and plot π_n, $n = 0,1,\ldots 5$ in one figure.

[2] Showing Legendre polynomials form an orthogonal system, while straightforward, is a good exercise in integration by parts! For a reasonably readable proof of completeness of Legendre polynomials in $L^2[-1,1]$ see for instance [2, p. 61].

See Matlab code, main script file (HW2_Q4.m) under label %%Part (a).



(b) Write a computer code that given a function $f \in L^2[-1,1]$, and a given approximation order p, computes the least squares approximation,

$$\varphi^*(x) = \sum_{j=0}^{p} \frac{(f,\pi_j)}{(\pi_j,\pi_j)}\pi_j(x).$$

Note that you know the analytic expression for (π_j,π_j) from (2). To evaluate $(f,\pi_j) = \int_{-1}^{1} f(x)\pi_j(x)\,dx$ you can use quadrature. Since we have not yet covered numerical integration, at this point you can use the routine `gauss_quad.m`, available in Moodle page for the course, which returns nodes $\{x_i\}_{i=1}^N$ and weights $\{w_i\}_{i=1}^N$ of an N-point Gauss–Legendre quadrature on $[-1,1]$; using this

$$\int_{-1}^{1} f(x)\pi_j(x)\,dx \approx \sum_{i=1}^{N} w_i f(x_i)\pi_j(x_i).$$

Use a large N to ensure results are accurate.[3]

Note that the main computational expense in real applications is the evaluation of the function f at the quadrature nodes (for example, in some applications, f might involve solving PDEs). Therefore, set up your implementation so that the function evaluations are done once at the beginning and then reuse the function evaluations $\{f(x_i)\}_{i=1}^N$, when computing each of the coefficients.

[3]Later in the course, we will discuss numerical integration in detail, and in particular, will learn about the degree of accuracy of such quadrature rules.

See Matlab code, function leastSq.m.
Other relevant functions – ip_fpi.m.

(c) Consider the following functions:

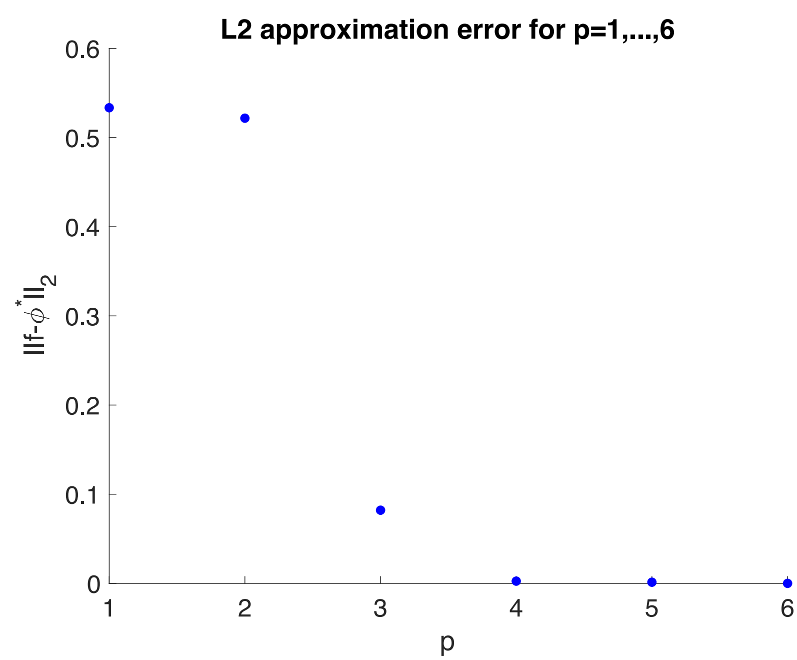
$$f(x) = e^x \sin(\pi x), \quad g(x) = 1/(1+x^2), \quad h(x) = |x| + x^2 \sin(\pi x) \qquad \text{for } x \in [-1, 1].$$

Use your routine from the previous part to compute the least squares approximating polynomials of order $p = 1, \ldots, 6$ for these functions. Provide both a table and a plot that for each p shows the approximation error $\|f - \varphi^*\|_2$. For each of the functions, provide a plot of the function with the approximations of order $j = 1, \ldots, 6$ in the same Figure. Also, report the expansion coefficients for $p = 6$ for each of the functions. Discuss the results briefly.

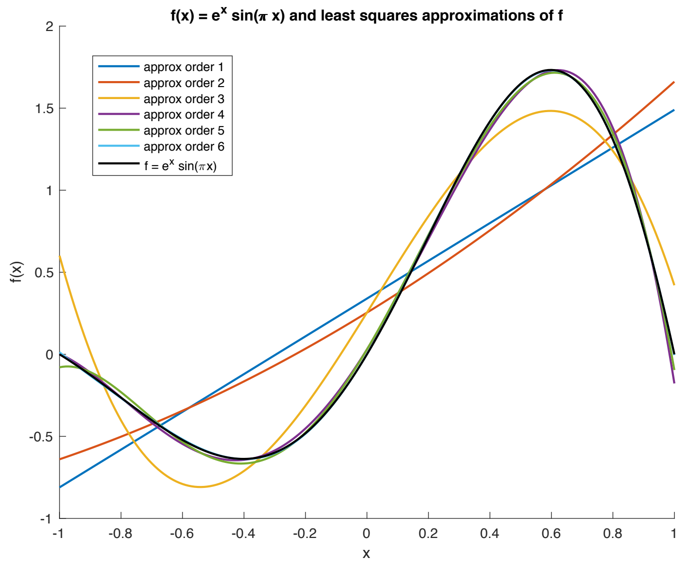
See Matlab code, main script file (HW2_Q4.m) under label %%Part (c).
Relevant functions – leastSq.m, ip_fpi.m, calcError.m, plotFandApproxs.m

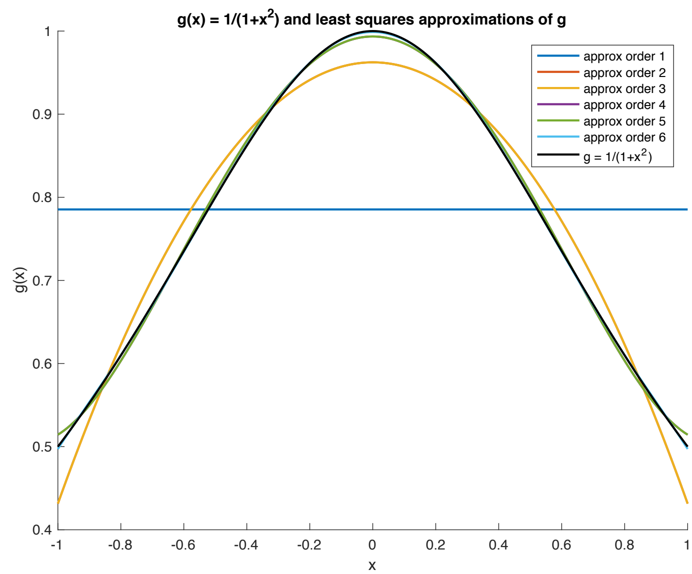
- Approximation error $\|f - \varphi^*\|_2$ for $p = 1, \ldots, 6$.

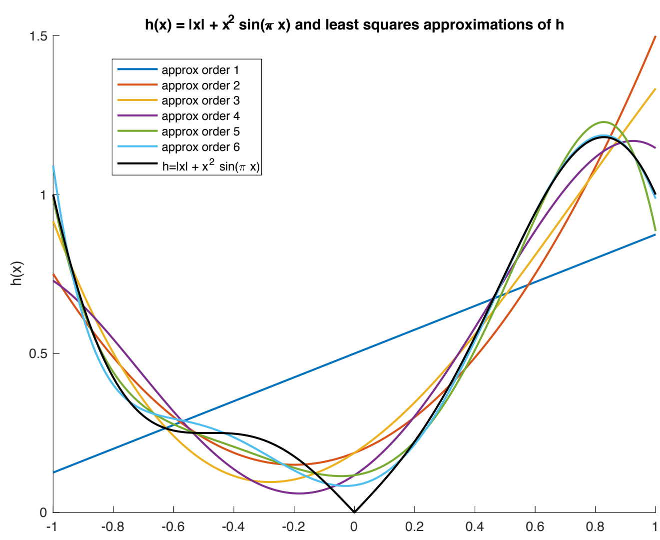
p	$\|f - \varphi^*\|_2$
1	0.533457242128192
2	0.521740530001147
3	0.082062453700881
4	0.002514335161159
5	0.001302378068193
6	0.000013580824437



- Plots of each function with approximations of order $j = 1, \ldots, 6$.







- Expansion coefficients for $p = 6$ for each of the functions.

  $f(x) = e^x \sin(\pi x)$:

j	0	1	2	3	4	5	6
$(f, \pi_j)/(\pi_j, \pi_j)$	0.3397	1.1505	0.1711	-1.2405	-0.5983	0.0816	0.0915

  $g(x) = 1/(1 + x^2)$

j	0	1	2	3	4	5	6
$(g, \pi_j)/(\pi_j, \pi_j)$	0.7854	0	-0.3540	0	0.0830	0	-0.0172

  $h(x) = |x| + x^2 \sin(\pi x)$

j	0	1	2	3	4	5	6
$(h, \pi_j)/(\pi_j, \pi_j)$	0.5000	0.3744	0.6249	-0.1657	-0.1874	-0.2607	0.1014

- Discuss the results briefly.

  The function $f(x) = e^x \sin(\pi x)$ is approximated well for sufficiently large n. For values of $p = 4, 5, 6$, the L_2 approximation error is less than 0.003.

  The function $g(x) = 1/(1+x)^2$ is also approximated well for sufficiently large n. For values of $p = 4, 5, 6$, the L_2 approximation error is less than 10^{-4}. Additionally, g is notable in that the polynomial approximation of degree ≤ 3 that minimizes the L_2 error is actually of degree 2 – and consequently is the same polynomial as the least squares approximation where $p = 2$. This is also true for the approxiations where $p = 4$ and $p = 5$ – both approximating polynomials are of degree 4.

  The L_2 error on the interval $[-1, 1]$ is minimized reasonably well for the function $h(x) = |x| + x^2 \sin(\pi x)$; however, the polynomials do a poor job of approximating the shape of the curve. In particular, the singularity at $x = 0$ is not well represented in the polynomials, which of course are continuous and have continuous derivatives. In this case, h would best be approximated by two distinct polynomials – one on the interval $[-1, 0]$ and another on the interval $[0, 1]$.

(d) Briefly explain how you would change the implementations to solve a least squares approximation problem as described above but with the interval $[-1, 1]$ replaced with any other finite interval $[a, b]$. You do not need to submit computer code that implements this modification.

To implement the problem on some interval $[a, b]$, I would need a weight function defined on the new interval. Additionally, new polynomials which are orthogonal on the new interval would need to be computed.
Since (f, π_j) would need to be calculated by integrating on the interval $[a, b]$, I would need to the use qrule.m function to return weights and x_is on the interval $[a, b]$ instead of $[-1, 1]$.
Also, I would need a new expression to compute (π_j, π_j).

5. [5 points] Consider the following discrete least squares approximation problem: given data $\big(x_i, f(x_i)\big)$, $i = 1, 2, \ldots, N$, find $\varphi^* \in \Phi_n = \text{span}\{\pi_1, \pi_2, \ldots, \pi_n\}$ that minimizes the error

$$E(\varphi) := \sum_{i=1}^{N} w_i |f(x_i) - \varphi(x_i)|^2, \quad \varphi \in \Phi_n,$$

where $w_i > 0$, $i = 1, \ldots, N$. Here we use the (discrete) inner product $(f, g) = \sum_{i=1}^{N} w_i f(x_i) g(x_i)$ and the norm $\|f\|_2 = (f, f)^{1/2}$. The set $\{\pi_j\}_{j=1}^{n}$ is linearly independent but not necssarily orthogonal. Assume $N > n$.

Consider the normal equations for the coefficients of the least squares approximation, as derived in the class, and show that it is equivalent to the following linear system:

$$\mathbf{K}^T\mathbf{W}\mathbf{K}\boldsymbol{c} = \mathbf{K}^T\mathbf{W}\boldsymbol{f}, \tag{3}$$

where $K_{ij} = \pi_j(x_i)$, and $f_i = f(x_i)$, $i = 1, \dots, N$, $j = 1, \dots, n$, and $\mathbf{W} \in \mathbb{R}^{N\times N}$ is a diagonal matrix with $w_1, \dots, w_N$ on its diagonal.

Want to show $\mathbf{A}\boldsymbol{c} = \boldsymbol{F}$ is equivalent to $\mathbf{K}^T\mathbf{W}\mathbf{K}\boldsymbol{c} = \mathbf{K}^T\mathbf{W}\boldsymbol{f}$.

- $\underline{(K^TW)_{ij}}$:

$$(K^TW)_{ij} = \pi_i(x_j)w_j$$

- $\underline{(K^TWK)_{ij}}$:

$$(K^TWK)_{ij} = \sum_{k=1}^{N} \pi_i(x_k)w_k\pi_j(x_k) = \sum_{k=1}^{N} w_k\pi_i(x_k)\pi_j(x_k) = (\pi_i, \pi_j) = A_{ij}$$

- $\underline{(K^TWf)_i}$:

$$(K^TWf)_i = \sum_{k=1}^{N} \pi_i(x_k)w_kf_k = \sum_{k=1}^{N} w_kf(x_k)\pi_i(x_k) = (f, \pi_i) = F_i$$

Since $(K^TWK)_{ij} = A_{ij}$ and $(K^TWf)_i = F_i$, $\mathbf{K}^T\mathbf{W}\mathbf{K} = \mathbf{A}$ and $\mathbf{K}^T\mathbf{W}\boldsymbol{f} = \boldsymbol{F}$.
Therefore $\mathbf{A}\boldsymbol{c} = \boldsymbol{F}$ is equivalent to $\mathbf{K}^T\mathbf{W}\mathbf{K}\boldsymbol{c} = \mathbf{K}^T\mathbf{W}\boldsymbol{f}$.

6. [10 points] Problem 20 at the end of Chapter 2 of the textbook [1, p. 123].

   Given the recursion relation $\pi_{k+1}(t) = (t-\alpha_k)\pi_k(t) - \beta_k\pi_{k-1}(t)$, $k = 0, 1, 2, \dots$ for the monic orthogonal polynomials $\{\pi_j\}_{j=0}^{\infty}$, and defining $\beta_0 = \int_a^b w(t)dt$ show that $\|\pi_k\|^2 = \beta_0\beta_1\beta_2\cdots\beta_k$, $k = 0, 1, 2, \dots$.

By theorem, the coefficients β_k are of the form

$$\beta_k = \frac{(\pi_k, \pi_k)}{(\pi_{k-1}, \pi_{k-1})} \qquad k = 1, 2, \dots.$$

And since $\pi_0 \equiv 1$, $\beta_0 = \int_a^b w(t)dt = (\pi_0, \pi_0)$.
Thus

$$\beta_0\beta_1\beta_2\cdots\beta_k = (\pi_0, \pi_0)\frac{(\pi_1, \pi_1)}{(\pi_0, \pi_0)}\frac{(\pi_2, \pi_2)}{(\pi_1, \pi_1)}\cdots\frac{(\pi_k, \pi_k)}{(\pi_{k-1}, \pi_{k-1})} = (\pi_k, \pi_k) = \|\pi_k\|^2.$$

References

[1] Gautschi, Walter. Numerical analysis. Springer Science & Business Media, 2011.

[2] Luenberger, David G. Optimization by vector space methods. John Wiley & Sons, 1997.