Problem Set 6

Problem 1. *Naïve curve fitting and its consequences*

Just like we were able to construct a line of best fit for a collection of data points, we can use the least-squares method to construct a "polynomial of best fit" instead. Unless care is taken, we are going to run into to very nasty problems rather quickly. The following exercises outlines some naïve approaches and some suggestions for how to fix the problems encountered with each.

(a) For any positive integer n, define the set

$$S_n = \{-n, -n+1, \ldots, -1, 0, 1, \ldots, n-1, n\} \subset \mathbb{R},$$

and take $f(x) = \sin(x)$. Use the model

$$y(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{2n} x^{2n} = \sum_{k=0}^{2n} a_k t^k$$

to create "parabola of best fit" for $f(x)$ on S_1. Repeat this to get a "quartic of best fit" on S_2 and a "sextic(?) of best fit" on S_3.

(b) Write a MATLAB script that computes the degree-$2n$ curve of best fit for $f(x)$ on S_n. The backslash operator $A\backslash b$ in MATLAB automatically produces least-squares solutions to $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ when $\vec{\mathbf{b}} \notin \operatorname{im} A$.

Plot your regression curves and $f(x)$ together on the same graph for the cases when $n = 1, 2, 3, 4, 5$. Make a conjecture about what you expect this regression curve to approach as $n \to \infty$?

(c) Try justifying your conjecture from part (b) by using your script to plot the degree-$2n$ regression curves for $n = 10, 20, 30$. What do you see in the plots of $y(x)$ compared with $f(x)$ as n gets larger? Does this prove or disprove your conjecture?

(d) We can try to fix the problem found in part (c) by using a different set of (smaller) sample points

$$T_n = \left\{ \frac{-n+j}{n} \right\}_{j=0}^{2n} \subset \mathbb{R}.$$

Append to your script from (b) another script that computes the degree-$2n$ curve of best fit for $f(x)$ on T_n instead and plots it with $f(x)$. Use this script for the cases when $n = 10, 20, 30$. Does this new script fix the problem in part (c)? Did we introduce any new problems with this new script?

(e) We can try to fix the problem found in part (d) by fixing the degree of our model to be 10, say. That is, let's fix our model as

$$y(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_{10} x^{10} = \sum_{k=0}^{10} a_k x^k.$$

Now append to your script yet another script that computes the degree-10 curve of best fit for $f(x)$ on T_n when $n > 10$ via the normal equations for the resulting linear system on the a_k variables (don't use MATLAB's backslash operation!). Use your script to create the degree-10 curve of best fit for $f(x)$ on T_n in the cases when $n = 20, 30, 40$. Does this new script fix the problem in part (d)? Did we introduce yet more problems with this third script?

(f) Suggest another change that we could make to our approximation scheme that you think would help prove your conjecture in part (b).

Please submit all three parts of your script.

Problem 2. *The uncanny sameness of finite-dimensional vector spaces*

One of the reasons we like finite-dimensional vector spaces so much is because of how similar they are to one another. Functions, matrices, signals... They all act like column vectors when we use them. They sure "look" and "feel" like column vectors, don't they? Let's formalize this a bit, shall we?

(a) Prove the following theorem:

Theorem 1. Let $(V, \mathbb{F})$ be a vector space with $\dim(V) = n < \infty$. Then $V \simeq \mathbb{R}^n$.

This shows that "all n-dimensional vector spaces are really the same."

(b) Show that this is false in the case where $\dim(V) = \infty$; that is, find two vector spaces V and W that both have infinite dimension but are not isomorphic. [*Hint*: You may use the fact that the integers $\mathbb{Z}$ and the real numbers $\mathbb{R}$ have different sizes, even though they are both infinite.]

Problem 3. *Sturm-Liouville, part II*

Recall the general Sturm-Liouville equation from the previous Problem Set:

$$-\frac{d}{dx}\left(p(x)\frac{dy}{dx}\right) + q(x)y = \lambda r(x)y,$$

where $p, q, r \in C^1[a, b]$ with $p(x), q(x), r(x) > 0$ for $a < x < b$ and $\lambda \in \mathbb{R}$ (note the leading minus sign now). For simplicity, we also stipulate that the *solutions* to this differential equation must satisfy the "pinned" *Dirichlet boundary condition*

$$y(a) = y(b) = 0.$$

Our goal for this exercise is to build up to the *weak formulation* of this differential equation. This allows us to frame the *solutions* to this differential equation as those functions that minimize a particular quadratic function. That is, we are going to show that this equation *always* has a *unique* solution satisfying the boundary conditions.

We first define the linear function $L\colon C_0^\infty[a, b] \to (C_0^\infty[a, b])^2$ as

$$L[u] = \begin{pmatrix} u' \\ u \end{pmatrix};$$

that is, L sends a smooth function (that is pinned at $x = a$ and $x = b$) to the vector containing its derivative and itself (in that order). If we take

$$\vec{\mathbf{f}} = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} \quad \text{and} \quad \vec{\mathbf{g}} = \begin{pmatrix} g' \\ g \end{pmatrix},$$

define the weighted inner product on $(C_0^\infty[a, b])^2$ as

$$\langle \vec{\mathbf{f}}, \vec{\mathbf{g}} \rangle = \int_a^b [p(x)f_1(x)g_1(x) + q(x)f_2(x)g_2(x)]\, dx.$$

(a) Show that the adjoint L^* of L with respect to this inner product is given by

$$L^* \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = -\frac{d(p(x)h_1(x))}{dx} + q(x)h_2(x)$$

$$= -p(x)h_1'(x) + p(x)h_1(x) + q(x)h_2(x).$$

(b) We know that $S = L^* \circ L$ is self-adjoint. Show that

$$S[u] = -\frac{d}{dx}\left(p(x)\frac{du}{dx}\right) + q(x)u(x)$$

so that we see the Sturm-Liouville equation as the linear equation

$$S[u] = \lambda r(x)u(x).$$

(c) Use the Minimization Theorem for linear functions to show that the solution to the Sturm-Liouville equation is given as the unique minimizer of the quadratic function

$$Q[u] = \frac{1}{2}\|L[u]\|^2 - \langle\langle \lambda r u, u\rangle\rangle,$$

where $\|\cdot\|$ is induced by $\langle\cdot,\cdot\rangle$, but $\langle\langle\cdot,\cdot\rangle\rangle$ is the standard L^2 inner product on $[a, b]$.