Homework 1

AMATH 563, Spring 2025

Due on Apr 14, 2024 at midnight.

Directions, Reminders and Policies

- You must upload two pdf files, one for theory questions and one for computational tasks to Gradescope.
- Make sure your solutions are well-written, complete, and readable for the theory question. I suggest you use Latex for your theory problems as well as your computational reports.
- For the computational tasks submit a pdf file typeset in Latex or any other typesetting software that can render equations appropriately. Please use the latex template provided on Canvas.
- I encourage collaborations and working with your colleagues to solve HW problems but you should only hand in your own work. We have a zero tolerance policy when it comes to academic misconduct and dishonesty including: Cheating; Falsification; Plagiarism; Engaging in prohibited behavior; Submitting the same work for separate courses without the permission of the instructor(s); Taking deliberate action to destroy or damage another person's academic work. **Such behavior will be reported to the UW Academic Misconduct office without warning**.

Theory Problems

Hint: read chapters 2 and 3 of Kreyszig.

1. Prove that $C([a,b])$ equipped with the $L^2([a,b])$ norm is not a Banach space.
2. If $(X_1, \|\cdot\|_1)$ and $(X_2, \|\cdot\|_2)$ are normed spaces, show that the (Cartesian) product space $X = X_1 \times X_2$ becomes a normed space with the norm $\|x\| = \max(\|x_1\|_1, \|x_2\|_2)$ where $x \in X$ is defined as the tuple $x = (x_1, x_2)$ with addition and scalar multiplication operations: $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$.
3. Show that the product (composition) of two linear operators, if it exists, is a linear operator.
4. Consider the functional $f(x) = \max_{t\in[a,b]} x(t)$ on $C([a,b])$ equipped with the sup norm. Is this functional linear? is it bounded?
5. Let X be a Banach space and denote its dual as X^*. Show that $\|\varphi\| : \varphi \mapsto \sup_{\|x\|=1} |\varphi(x)|$ is a norm on X^*.
6. Prove the Schwartz inequality on inner product spaces: $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$ for all $x, y \in X$, where equality holds if and only if x, y are linearly dependent.

Computational Problems

Solve the computational problem below. Perform the required tasks and hand in a report of a maximum of four pages outlining your methodology, results, and findings. Use the computational report template provided on Canvas. You don't have to use Latex but it is highly recommended. Make sure that your figures and tables are readable, they are of good size, and labeled appropriately. Part of your grade will be assigned to the general tidiness and style of the report. **You don't need to hand in your code.**

Consider the Levy function in 1D

$$f(x) = \sin(\pi w)^2 + (w-1)^2\Big(1 + \sin(2\pi w)^2\Big), \quad w = 1 + \frac{10x - 1}{4}, \quad x \in [-1, 1],$$

a plot of this function is shown in Figure 1 with the blue solid line. Now consider a uniform grid of points $\{x_k\}_{k=1}^K \in [-1, 1]$ with $K = 50$ and suppose we have the observations

$$y_k = f(x_k) + \xi_k, \qquad \xi_k \sim N(0, \sigma^2), \quad k = 0, \dots K-1,$$

where $\sigma > 0$ is the standard deviation of the observation noise ξ_k. Write $\boldsymbol{y} \in \mathbb{R}^K$ for the vectors of y_k values. The red dots in Figure 1 show an instance of these observations. *Your goal in this problem is to find an approximation to f using the limited and noisy observations $\boldsymbol{y}$.*

Since we don't know what f is we have to introduce a model. Consider functions of the form

$$g(x) = \sum_{n=0}^{N-1} c_n \psi_n(x),$$

for integer $N \geq 1$, coefficients $c_n \in \mathbb{R}$, and a collection/dictionary of functions ψ_n. Different choices of the ψ_n yield different models for the function f. We will approximate f over the dictionary $\{\psi_n\}_{n=0}^{N-1}$ by solving the optimization problem

$$\widehat{f} := \arg\min_{\boldsymbol{c}} \frac{1}{2} \sum_{k=0}^{K-1} |g(x_k) - y_k|^2 + \frac{\lambda}{2}\|\boldsymbol{c}\|_2^2, \quad \text{subject to} \quad g(x) = \sum_{n=0}^{N-1} c_n \psi_n(x). \tag{1}$$

where we wrote $\boldsymbol{c} \in \mathbb{R}^N$ for the vector of the c_n coefficients and $\lambda > 0$ is a regularization parameter.

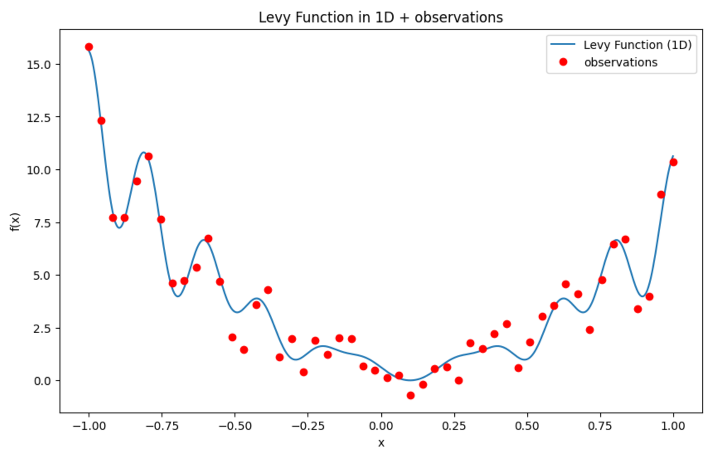

Figure 1 Plot of the 1D Levy function along with 50 observation points generated by the setup script.

1. Rewrite (1) as a least squares problem

$$\widehat{f} = \sum_{n=0}^{N-1} \widehat{c}_n \psi_n, \qquad \widehat{c} := \arg\min_{\boldsymbol{c} \in \mathbb{R}^N} \frac{1}{2}\|\Psi \boldsymbol{c} - \boldsymbol{y}\|^2 + \frac{\lambda}{2}\|\boldsymbol{c}\|^2,$$

   for a matrix $\Psi \in \mathbb{R}^{K \times N}$ and write down an analytic expression for the minimizer $\widehat{c}$ in terms of $\Psi, \lambda, \boldsymbol{y}$.
2. Write a script to compute $\widehat{f}$ over the following dictionaries. Describe your algorithm of choice in the report:
   - (Cosine series) $\psi_n(x) = \cos(n\pi x)$
   - (Monomials) $\psi_n(x) = x^n$
   - (RBF basis) $\psi_n(x) = \exp\left(-\frac{\|x - z_n\|^2}{2\ell^2}\right)$ for $\ell = 0.2$ and $z_n = -1 + 2n/N$.
3. Compute the $\widehat{f}$ for each of the above models for $N = 10, 50, 100$ and for $\lambda = 10^{-10}, 10^{-2}, 1$ and present a plot of $\widehat{f}$ alongside f. You may overlay multiple plots for each model in the same figure. Be sure to use legends as needed if you do.
4. Continue to investigate the effect λ and how it interacts with N and choice of the dictionary. You can also consider other dictionaries. Report your findings.