APMA 2070 & ENGN 2912V Deep Learning for Scientists and Engineers
Homework 03
Due Date: 03-05-2024, 11:59 pm (E.T.)

Lecture 1.3

1. Write a TensorFlow code to approximate the oscillatory function using tanh and ReLU activation functions, and compare your results for 40 and 80 data points that are (a) equi-spaced and (b) randomly sampled from a uniform distribution.

$$\begin{aligned} y &= 5 + \sum_{k=1}^{6} \sin(kx), \quad x < 0 \\ y &= \cos(10x), \ x \geq 0 \end{aligned} \tag{1}$$

2. Write TensorFlow and PyTorch code to construct the Laplacian operator in 3D. Test on the following example

   (a) $f(x, y, z) = \sqrt{xy} + z$

   (b) $f(x, y, z) = \sqrt{3x^2 + y^2 + 5z^2}$

   (c) $f(x, y, z) = x^7 y^5$

   (d) $f(x, y, z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$

3. Write TensorFlow/PyTorch code to learn the oscillatory function

$$f(x) = \begin{cases} 5 + \sum_{k=1}^{4} \sin(kx), & x \in [-\pi, 0) \\ \cos(10x), & x \in [0, \pi] \end{cases} \tag{2}$$

   using a shallow ReLU network (i.e., one hidden layer) and 200 uniform data points in $[-\pi, \pi]$ (i.e., training dataset).

   Use networks of widths $\{10, 30, 100, 300, 1000\}$, and compute the L^2 relative errors of these networks on a testing dataset consisting of at least 500 equi-spaced data points in $[-\pi, \pi]$. You can use MSE loss function for training the neural network model. Make sure that the network is trained well by choosing the appropriate learning rate and adequate number of epochs until convergence. Run your code at least 10 times from different random weight initializations (refered as *seed* in python) and compute the mean and standard deviation of the prediction errors for each width. Plot the error versus the network width, and discuss what you observe.

Hint: The relative L^2 error between the network $\mathcal{N}(x)$ and $f(x)$ is defined as

$$\frac{\|\mathcal{N} - f\|_2}{\|f\|_2},$$

where the L^2-norm $\|\cdot\|_2$ for a function $g(x)$ ($x \in \Omega$) is defined as

$$\|g\|_2 = \left(\int_\Omega g^2(x)dx\right)^{\frac{1}{2}}.$$

Hence, the relative L^2 error can be approximated by

$$\frac{\|\mathcal{N} - f\|_2}{\|f\|_2} \approx \frac{\sqrt{\sum_{i=1}^n \left(\mathcal{N}(x_i) - f(x_i)\right)^2}}{\sqrt{\sum_{i=1}^n f^2(x_i)}},$$

where $\{x_i\}_{i=1}^n$ are n equi-spaced points in Ω; n should be large enough to make the approximation accurate.

(a) Repeat Problem (3) using two hidden layers.

Compare the new results and the results in Problem (3). Discuss your observations.

Hint: Shallow networks vs Deep networks.

(b) Repeat Problem (3) using 20 data points in $[-\pi, \pi]$.

Compare the new results and the results in Problem (3). Discuss your observations.

Hint: Estimation error.