

# ECE 512 – Topics in Data Science

## Homework 1

Dror Baron and Hangjin Liu; Fall 2023

Due: August 30, 2023

### Administrative instructions:

1. For any clarification or doubts, the TA Hangjin Liu (hliu25 AT ncsu DOT edu) is in charge of homeworks and projects. She should be your first point of contact on homework- and project-related issues.
2. The homework can be submitted individually, in pairs, or triples.
3. You should submit electronically through Moodle by midnight the day that the homework is due.
4. Please justify your answers carefully.

### 1. Polynomial curve fitting:

Recall the comprehensive example on polynomial curve fitting (LecturesPPT1.pdf; slides 22-32).

- a. Define 2-3 nonlinear functions. For each function, create test points and show that when we increase the number of observations  $N$ , and keep the model order  $M$  fixed, the polynomial function being learned becomes closer to the true nonlinear function. Investigate this dependence both qualitatively (plot how the learned functions approach the true ones) and quantitatively (define some error function, for example square error, and see how the error goes down as  $N$  is increased).
- b. Investigate the dependence on the noise variance. In particular, keep  $M$  and  $N$  fixed, and show qualitatively and quantitatively how reducing the noise improves the learned function.
- c. Does curve fitting actually “work?” That is, if we learn a polynomial function using training data, will it fit the test data well? Design an experiment that aims to show this using your simulation setup, perform these tests, and discuss/show your findings. This will likely involve (i) dividing the data into training / test; (ii) training the model; and (iii) showing how far the predicted value differs from the actual value of the test point.

### 2. Probability:

Suppose that, on a given day, a computer server crashes (C) with probability  $\frac{1}{12}$  and resets (R) with probability  $\frac{1}{9}$ ; the probability that it both resets and crashes on a given day is  $\frac{1}{36}$ .

- a. Determine the probability that it (i) crashes without resetting, and (ii) resets without crashing.
- b. Determine the probability that it (i) crashes or resets (or both), and (ii) neither crashes nor resets.
- c. Determine the conditional probability that it (i) resets given that it crashes, and (ii) resets given that it doesn't crash.

### 3. Probability:

Suppose  $f_X(x) = x(u(x) - u(x - 1)) + a\delta(x) + 2a\delta(x + 1)$ , where  $u$  denotes a step function and  $\delta$  a delta impulse.

- a. Determine  $a$ . Use this value of  $a$  in the remainder of this problem.
- b. Determine  $F_X(x)$ .