CS57300 Spring 2018 Homework — Homework 0 — Last update: December 19, 2017

https://www.cs.purdue.edu/homes/ribeirob/courses/Spring2018/index.html Due: January 8th 8:00am, 2018

Instructions and Policy: Each student should write up their own solutions independently, no copying of any form is allowed. You MUST to indicate the names of the people you discussed a problem with; ideally you should discuss with no more than two other people.
YOU MUST INCLUDE YOUR NAME IN THE HOMEWORK
You need to submit your answer in PDF. LaTeX is typesetting is encouraged but not required. Please write clearly and concisely - clarity and brevity will be rewarded. Refer to known facts as necessary.

Q0 (0pts correct answer, -1,000pts incorrect answer: (0,-1,000) pts): A correct answer to the following questions is worth 0pts. An incorrect answer is worth -1,000pts, which carries over to other homeworks and exams, and can result in an F grade in the course.

(1) Student interaction with other students / individuals:

(a) I have copied part of my homework from another student or another person (plagiarism).

(b) Yes, I discussed the homework with another person but came up with my own answers. Their name(s) is (are) --

(c) No, I did not discuss the homework with anyone

(2) On using online resources:

(a) I have copied one of my answers directly from a website (plagiarism).

(b) I have used online resources to help me answer this question, but I came up with my own answers (you are allowed to use online resources as long as the answer is your own). Here is a list of the websites I have used in this homework:

(c) I have not used any online resources except the ones provided in the course website.

Q1 (4 pts):
Let $(X_1, \ldots, X_n)$, $n \geq 2$, where $X_i \sim Normal(0,1)$, $i = 1, \ldots, n$, are i.i.d. random variables. Answer the following questions. You must be able to use Python 3 for your code for any coding questions.

(a) Consider $n = 2$, (X_1, X_2) described above. Let $Y_1 = X_1 + 2X_2$ and $Y_2 = 2X_1 + X_2$. What is the distribution of (Y_1, Y_2)? What are the parameters of this distribution?

(b) Use matplotlib to plot a scatter plot with 100 samples of (X_1, X_2) within the 2D square (-2,-2), (2,2). Note: Some of the 100 samples might fall outside the square (ignore them).

(c) Empirically compute the probability that an observation of (X_1, X_2) falls within a circle of radius 0.5 centered at (0,0).

(d) Draw 100 observations from (X_1, X_2, X_3) and compute the empirical probability that an observation falls within a sphere of radius 0.5 centered at (0,0,0).

(e) Repeat the above exercise for $n = 1000$.

(f) Let $n \to \infty$. What is the probability that an observation falls within a hypersphere of radius 0.5 centered at the origin $(0, 0, \ldots)$? Prove it.

Q2 (3 pts): Probability, inference, and Bayes theorem.

(a) Prove the conditional version of Bayes rule:

$$P(B|A,C) = \frac{P(A|B,C)P(B|C)}{P(A|C)}$$

(b) If $X_1, \ldots, X_n$ are independent and identically normally distributed random variables, $X_i \sim N(\mu, \sigma^2)$. Estimate μ and σ via maximum likelihood in closed form.

(c) Repeat item (b) but now write a maximum 15-line Python 3 code using numpy to estimate μ and σ via maximum likelihood using gradient ascent.

(d) Suppose we do not know μ and let $\mu \sim \text{Normal}(\mu_0, \sigma_0)$ be the prior distribution of μ, where $\sigma_0 > 0$ and μ_0 are known constants. Assume $\sigma > 0$ is also a known constant. Prove that the posterior distribution $P[\mu|X_1, \ldots, X_n, \sigma]$ is

$$\mu \mid X_1, \ldots, X_n \sim \text{Normal}\left(\left(\frac{\mu_0}{\sigma_0^2} + \frac{\sum_{i=1}^n X_i}{\sigma^2}\right)\Big/\left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right), \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\right)^{-1}\right)$$

Q3 (2 pts): Probability and conditional probability.

(a) The Internet is a wonderful source of information about symptoms of rare diseases. Are you sneezing? It could be the West Nile virus! The West Nile virus (WNV) infected approximately 2,000 people in the United States last year[1]. Sheldon, your hypochondriac friend, is sneezing and heard about the West Nile virus on Twitter. He demands a test for the West Nile virus, why not? The test correctly identifies the presence of WNV in 95% of cases and only gives false positives in 1/10,000 cases. Unfortunately, the test came back positive for West Nile virus and Sheldon is very concerned. Assume that in the United States there are 300 million people susceptible to WNV.

(i) What is the probability that Sheldon has WNV?

(ii) The WNV virus is fatal in 4% of the cases. What is the probability that Sheldon will die this year? Assume a fatality rate of any cause (car accident, etc.) of 0.2% that is independent of whether or not Sheldon has WNV.

(b) Alice and Bob are playing a simple dice game. Each rolls one dice and the one with higher number wins. If the numbers are the same, they roll again. If Alice just won, what is the probability that she just rolled a "3"?

[1] Source: the "always-reliable" Wikipedia

Q4 (3 pts): (Python / numpy coding).

(a) Write a Python 3 script with at most 15 lines of code that reads a UTF-8 text file of name “matrix.csv” with format.

```
10,  100, winter
8 , 10, summer
0 , 0, spring
2 , 5, fall
18 , 4, summer
...
```

and creates a numpy matrix `A[i,j] = season`, where i is the number if the first column, j is the number if the second column, and `season` is encoded as winter = -1, summer=1, and fall=spring=0.

(b) Using the above matrix, show a 1-line Python 3 code that selects a submatrix with the 2nd and 3rd rows of the matrix, and the 10th and 11th columns (note that python array/matrix indices should start with 0).

(c) Let u and v be two numpy vectors both with dimensions $n \times 1$ (n-dimensional column vectors). Show how to compute the inner product between u and v in one operation using numpy (no python loops are allowed).

Q5 (2 pts): (Linear Algebra)
Let A be a 0-1 matrix (a matrix whose elements are 0 or 1) representing a graph (i.e., A is an adjacency matrix). Let A^T be the transpose of matrix A. Let $A = U\Sigma V^T$ be the singular value decomposition (SVD) of A. Describe how to obtain the SVD decomposition of $B = A^T A$ and $C = AA^T$ without having to redo the decomposition. Write down the necessary steps needed to show that no recalculation is necessary.