CS4786/5786: Machine Learning for Data Science, Fall 2017
08/22/2017: Assignment 0: Diagnostic

This assignment is more of a diagnostic test covering prerequisite concepts for this class. We want you to see feedback we'll provide regarding your level of preparation for the class, but we don't want you to stress out about being "judged". So, for this assignment, which should take *at most* two hours, we'll give credit on an all-or-none basis, as follows. If you turn the assignment in on time with good-faith attempts to show how you approached each question (*whether or not* you arrive at a correct answer or even get to an answer at all), you'll get full credit. That's it!

The assignment should be done *individually*. Don't discuss the problem set with other people. You may consult textbooks and other offline or online sources, but you should mention them and state what portions of your work they influenced.

Submit your solutions via the google form by **August 29th, 2017**. Some of the questions in this assignment are recycled from the first time the course was taught by Prof. Lee and me. I would like to thank Prof. Lee for her help in making the questions.

Academic integrity policy We distinguish between "merely" violating the rules for a given assignment and violating *academic integrity*. To violate the latter is to commit *fraud* by claiming credit for someone else's work. For this assignment, an example of the former would be getting an answer from person X but stating in your homework that X was the source of that particular answer. You would cross the line into fraud if you did not mention X. The worst-case outcome for the former is a grade penalty; the worst-case scenario in the latter is academic-integrity hearing procedures.

The way to avoid violating academic integrity is to always document any portions of work you submit that are due to or influenced by other sources, even if those sources weren't permitted by the rules.[1]

Q1 (Linear Algebra). Is it possible to have two n-dimensional vectors $\overrightarrow{x}$ and $\overrightarrow{y}$ to be orthogonal, i.e., $\overrightarrow{x} \cdot \overrightarrow{y} = 0$. And yet, have a $n \times n$ matrix A such that the vectors $A\overrightarrow{x}$ and $A\overrightarrow{y}$ are **not** orthogonal? Prove your answer, showing all work.

Incidentally, this question relates to *dimensionality reduction*, the first topic of the class.

Q2 (Eigen Vectors and Eigen Values). Consider any $n \times n$ square, symmetric matrix A (i.e. $A = A^\top$) such that the diagonal element of each row of this matrix is the negative sum of the remaining elements of that row. As an example consider the following matrix:

$$A = \begin{bmatrix} -4 & 1 & 3 \\ 1 & -8 & 7 \\ 3 & 7 & -10 \end{bmatrix}$$

[1] We make an exception for sources that can be taken for granted in the instructional setting, namely, the course materials. To minimize documentation effort, we also do not expect you to credit the course staff for ideas you get from them, although it's nice to do so anyway.

(PROBLEM(S) CONTINUE ON NEXT PAGE)

Note that for the above matrix, for any row, the diagonal element is the negative sum of the remaining elements of that row. Show that for any such $n \times n$ matrix A, the all ones vector $\overrightarrow{\mathbf{1}}$ of length n, (in the above example for instance the vector $\overrightarrow{\mathbf{1}} = [1,1,1]^\top$) is such that,

$$A\overrightarrow{\mathbf{1}} = 0$$

Effectively you are proving that the all ones vector (scaled appropriately) is an eigen vector for any such matrix with corresponding eigenvalue of 0. If you are not comfortable writing a formal proof, sketch your reasoning or explanation for this. (Hint: Think about why its true in the above example first.)

Q3 (Conditional Probability[2]). Ms. Y lives next to a house with a big lawn. On any given day, the probability that it rains in Ms. Y's neighborhood is 0.1. The probability that the sprinkler is turned on next door is 0.3. When it rains, the probability that Ms. Y wears a poncho is 0.6. When the sprinkler is on, the probability that Ms. Y wears a poncho is 0.2. She never wears a poncho unless it either rained or the sprinkler was on.

You meet Ms. Y wearing a poncho. Is it more likely that the sprinkler was turned on that day or that it rained in Ms. Y's neighborhood that day? Show your work.

Q4 (Bayes Rule and Marginalization). Let A, B and C be binary random variables (taking values in $\{0,1\}$). You are given the joint distribution over A, B, C as

A	B	C	P(A,B,C)
0	0	0	0.05
0	0	1	0.1
0	1	0	0.15
1	0	0	0.2
0	1	1	0.05
1	0	1	0.25
1	1	0	0.15
1	1	1	0.05

Write down the value of $P(A = 1|B = 0)$. Explain how you obtained this answer. (Hint: use Bayes rule)

[2] Problem inspired by a classic example regarding Bayes Nets. (You don't need to know what a Bayes Net is.)