DSC 40A - Homework 5

Due: Sunday, May 15, 2022 at 11:59pm PDT

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Homeworks are due to Gradescope by 11:59pm PDT on Sunday.

Homework will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should **always explain and justify** your conclusions, using sound reasoning. Your goal should be to convince the reader of your assertions. If a question does not require explanation, it will be explicitly stated.

Homeworks should be written up and turned in by each student individually. You may talk to other students in the class about the problems and discuss solution strategies, but you should not share any written communication and you should not check answers with classmates. You can tell someone how to do a homework problem, but you cannot show them how to do it.

This policy also means that you **should not post or answer homework-related questions on Piazza**, which is a written medium. This includes private posts to instructors. Instead, when you need help with a homework question, talk to a classmate or an instructor in their office hours.

For each problem you submit, you should **cite your sources** by including a list of names of other students with whom you discussed the problem. Instructors do not need to be cited. The point value of each problem or sub-problem is indicated by the number of avocados shown.

Problem 1. One Dimensional Linear Regression Done Two Ways

In the class, we discussed that we can actually solve the one dimensional linear regression problem in two ways: with the slope and intercept formula that we discussed before the midterm or with the normal equation. Let's check that they lead to the same solution.

In particular, answer the following two questions.

a) For a one dimensional regression problem, suppose that the data points are $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, what is the design matrix X and the response vector $\vec{y}$? Define the weight vector $\vec{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$. What is the normal equation in terms of $x_1, \ldots, x_n$, $y_1, \ldots, y_n$, and w_1, w_2?

b) Note that your normal equation can now be written as a system of two linear equations in w_1 and w_2. Simplify the normal equation and check that one of the two equations corresponds to the intercept formula for w_0. Plugging this expression for w_0 into the other equation and solve for w_1.

Problem 2. Linear Regression and Regularization

Suppose we are to learn the following linear relationship:

$$y = w_0 + w_1 x_1 + w_2 x_2.$$

Let $\vec{w} = (w_0, w_1, w_2)^T$ be the vector of parameters. Suppose we have the following data:

x_1	1	2	3	4
x_2	7	14	21	28
y	10	20	30	40

a) For this data set, what is the design matrix X and what is the rank of the matrix $X^T X$?

b) Can you give a unique solution $\vec{w}$ to the normal equations? If so, find this vector $\overrightarrow{w}$. If not, explain why not.

Now remember that vector(s) $\vec{w}$ that satisfy the normal equations also minimize the squared risk:

$$R_{sq}(\overrightarrow{w}) = \frac{1}{n}\|\overrightarrow{y} - X\overrightarrow{w}\|^2.$$

Let's now consider a new risk function, inspired by the squared risk, but with an additional term $\lambda\|\overrightarrow{w}\|^2$, where $\lambda > 0$ is a free parameter. We'll call this new risk function $R_\lambda(\vec{w})$:

$$\begin{aligned} R_\lambda(\overrightarrow{w}) &= R_{sq}(\overrightarrow{w}) + \lambda\|\overrightarrow{w}\|^2 \\ &= \frac{1}{n}\|\overrightarrow{y} - X\overrightarrow{w}\|^2 + \lambda\|\overrightarrow{w}\|^2. \end{aligned}$$

We'll find the vector(s) $\vec{w}$ that minimize this new risk function, and we'll see whether the addition of the $\lambda\|\overrightarrow{w}\|^2$ term makes a difference in the number of vectors $\vec{w}$ that minimize the risk.

c) Calculate the gradient of $R_\lambda(\overrightarrow{w})$ with respect to $\vec{w}$.

Hint: You'll need to use the results of your last homework.

d) Which matrix needs to have full rank for $R_\lambda(\overrightarrow{w})$ to have a unique minimizer $\vec{w}$? For the particular data set of size $n = 4$ given here, and assuming $\lambda > 0$ is fixed, is there a unique minimizer of $R_\lambda(\overrightarrow{w})$? If so, find this unique minimizer $\vec{w}$, which may depend on λ. If not, explain why not.

e) Notice that as $\lambda \to 0$, $R_\lambda(\vec{w})$ approaches $R_{sq}(\vec{w})$. If we let $\lambda \to 0$ in part (d), do we obtain a unique solution for $\overrightarrow{w}$? Compare with part (b) and explain why this is the case.

Problem 3. Jackpot Dilemma

You are invited on a TV program where you have the opportunity to win a jackpot worth eight million dollars. You enter the stage and there are four boxes on the table. Only one of the four boxes has the jackpot money. Without other information, you can assume that the probability of each of box having the prize is uniform.

a) The host first asks that you make a choice among the four boxes. You simply choose a random box and remove it from the table. For the box that you choose, what is the probability of you hitting the jackpot? What is the sample space and what are the probabilities associated with the elements in the sample space?

b) After a while, a magician comes to the stage and opens up one of the three boxes left on the table (not chosen by you). It turns out to be empty. The host then says: "Here is your million dollar chance. You get to either keep the box you are holding or change your choice to one of the unopened boxes on the table."

Now what is the probability of the box in your hands containing the prize? And what is the probability of one of the unopened boxes on the table containing the prize? What does the host mean when he says, quite literally, "your million dollar chance"? Should you change the box you've chosen?

Explain your answer using conditional probability.

Problem 4. Rule of Probabilities

a) The multiplication rule for two events says

$$P(A \cap B) = P(A) \cdot P(B|A).$$

Use the multiplication rule for two events to prove the multiplication rule for three events:

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|(A \cap B)).$$

b) The general addition rule for any two events says:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Use the general addition rule for two events to prove the general addition rule for three events:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(C \cap B) + P(A \cap B \cap C).$$

c) Draw a Venn diagram to illustrate the general addition rule for three events and explain the general addition rule for three events in terms of your diagram.

Problem 5. House of Cards

A standard deck of card contains 52 cards. There are 13 cards in each of 4 suits (hearts ♡, spades ♠, diamonds ♢, and clubs ♣.) Within a suit, the 13 cards each have a different rank. In ascending order, these ranks are 2, 3, 4, 5, 6, 7, 8, 9, 10, Jack, Queen, King, Ace.

Suppose you are playing a four-player card game using two regular decks of cards. Each player will be dealt 26 cards. In this game, two identical cards, with the same rank and same suit, are called a pair. A pair can only be outranked (beaten) by a pair of higher ranked cards within the same suit. For example, a pair of Queens of Hearts can be outranked by a pair of Kings of Hearts, which in turn can only be outranked by a pair of Aces of Hearts.

a) Suppose the first deck of cards is randomly shuffled and dealt out to the four players. Then the second deck of cards is randomly shuffled and dealt out to the four players. What is the probability that you get a pair of Kings of Hearts?

b) Suppose the first deck of cards is randomly shuffled and dealt out to the four players. Then the second deck of cards is randomly shuffled and dealt out to the four players. What is the probability that some other player gets a pair of Aces of Hearts?

c) Suppose the first deck of cards is randomly shuffled and dealt out to the four players. Then the second deck of cards is randomly shuffled and dealt out to the four players. What is the probability that you have no pairs?

d) Suppose the first deck of cards is randomly shuffled and dealt out to the four players. Then the second deck of cards is randomly shuffled and dealt out to the four players. You are dealt two Kings of Hearts. What is the probability that some other player has a pair of Aces of Hearts?

e) Suppose the first deck of cards is randomly shuffled and dealt out to the four players. Then the second deck of cards is randomly shuffled and dealt out to the four players. You are dealt two Kings of Hearts and you do not have any Aces of Hearts. What is the probability that some other player has a pair of Aces of Hearts?

f) Suppose both decks of cards are first shuffled together and then dealt out to the four players. What is the probability that you get a pair of Kings of Hearts?

g) Compare your answers to parts (a) and (f). Are you more or less likely to get a pair of Kings of Hearts when you shuffle the decks together, or does it not matter? Explain intuitively why this is the case.