DSC 40A - Redemption Homework

Due: Sunday, June 5, 2022 at 11:59pm PDT

This is an optional assignment that you can do to replace one homework score. 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. Different Loss Functions

Consider three new loss functions,

$$\boxed{\mathrm{L}_1(h;y) = 2|h-y|-1} \qquad \boxed{\mathrm{L}_2(h;y) = \begin{cases} 2|h-y|-1, & \text{for } h-y>1 \\ (h-y)^2, & \text{for } h-y \le 1 \end{cases}} \qquad \boxed{\mathrm{L}_3(h;y) = \begin{cases} 2|h-y|-1, & \text{for } h-y<1 \\ (h-y)^2, & \text{for } h-y \ge 1 \end{cases}}.$$

a) Fix an arbitrary value of y. On three separate axes, **draw the graphs** of $L_1(h;y), L_2(h;y)$, and $L_3(h;y)$ as functions of h (label your choice of y on the graph).

Suppose we fix a true parameter θ to be a random number from 0 to 10, and generate `=10 } data poits` uniformly at random in an interval of width 10 centered at θ:

```
n = 10
theta = np.random.uniform(0.0,10.0)
y = np.random.uniform(theta-5.0, theta+5.0, n)
```

Now we can make various predictions for the center of the interval by minimizing each of the empirical risk functions

$$\boxed{\mathrm{R}_1(h) = \frac{1}{n}\sum_{i=1}^{n} L_1(h;y_i)} \qquad \boxed{\mathrm{R}_2(h) = \frac{1}{n}\sum_{i=1}^{n} L_2(h;y_i)} \qquad \boxed{\mathrm{R}_3(h) = \frac{1}{n}\sum_{i=1}^{n} L_3(h;y_i)}$$

using the gradient descent algorithm. Suppose h_1, h_2, h_3 are the minimizers achieved by gradient descent for R_1, R_2, R_3, respectively.

b) **Compare each of h_1, h_2, h_3 to θ** and say whether you expect the prediction to be **smaller than, greater than, or about the same** as θ. For example, do you expect to have $h_1 > \theta$, $h_1 < \theta$, or $h_1 \approx \theta$? Justify your answers based on the graphs you drew above.

Problem 2. Vector Derivative

If $\overrightarrow{w} \in \mathbb{R}^d$, compute the gradient with respect to $\overrightarrow{w}$ of the length of $\overrightarrow{w}$:

$$\frac{d}{d\overrightarrow{w}} \|\overrightarrow{w}\|.$$

Problem 3. Linear Regression

One of your classmates asked the following question about multivariate linear regression. Suppose we are to learn the following linear relationship:

$$y = w_0 + w_1 x^{(1)} + w_2 x^{(2)}. \tag{1}$$

What if instead of performing linear regression jointly over $x^{(1)}$ and $x^{(2)}$, we perform linear regression on each of the variables and combine them. That is, we first learn the linear relationship of

$$y = w_0^{(1)} + w_1 x^{(1)}. \tag{2}$$

We then learn the relationship of

$$y = w_0^{(2)} + w_2 x^{(2)}. \tag{3}$$

Can we recover the original parameters w_0, w_1, and w_2 using relationships in equations (2) and (3)? In particular, can we somehow combine $w_0^{(1)}$ and $w_0^{(2)}$ in equations (2) and (3) to obtain w_0 in equation (1)?

Suppose we collect n data points: $\left\{(x_1^{(1)}, x_1^{(2)}, y_1), \ldots, (x_n^{(1)}, x_n^{(2)}, y_n)\right\}$. Answer the following questions.

a) Assume that $\sum_{i=1}^n \left(x_i^{(1)} x_i^{(2)}\right) = \sum_{i=1}^n \left(\bar{x}^{(1)} \bar{x}^{(2)}\right)$, where $\bar{x}^{(1)} = \frac{1}{n}\sum_{i=1}^n x_i^{(1)}$ and $\bar{x}^{(2)} = \frac{1}{n}\sum_{i=1}^n x_i^{(2)}$. Could you obtain w_0 from combining $w_0^{(1)}$ and $w_0^{(2)}$ and recover equation (1) computed with multivariate linear regression? If not, give a counter example.

b) Now let's parse the condition of $\sum_{i=1}^n \left(x_i^{(1)} - \bar{x}^{(1)}\right)\left(x_i^{(2)} - \bar{x}^{(2)}\right) = 0$. In terms of the probability theorem, guess what does it stand for (you simply need to provide an intuition)?

Problem 4. Probability Theory

Let A and B be two independent events in the sample space S. Show that $\bar{A}$ and $\bar{B}$ must be independent of one another. Make sure your proof still works in the special case that one or more of A, B, $\bar{A}$, $\bar{B}$ has probability zero.

Hint: It helps to draw a Venn diagram.

Problem 5. Coping with Failure

When someone is faced with failure to an opponent, there are three possible states of mind: denial, fury, and acceptance. Assume that a person can only be in one state of mind at a certain point in time.

When in denial, there is a 5% chance they will congratulate their opponent. When in fury, there is a 10% chance they will congratulate their opponent. When in acceptance, there is a 90% chance they will congratulate their opponents.

Suppose that at a certain point in time, the failed party congratulated their opponent. What can we deduce about their state of mind? In this problem, we'll see that our answer depends upon the failed person's natural state of mind when they typically fail. For example, someone who is often accepting of their failures is more likely to be in a state of acceptance than someone who is rarely accepting of their failures, even if we know they both congratulated thier opponents.

a) Assume that normally, the failed party is equally likely to be in any of the three states when they fail. Knowing that they congratulated their opponent, what is the probability that the failed party is in denial? And what's the probability they are in fury and in acceptance, respectively?

b) Suppose that you know the failed party really well, and that when they fail, they are in denial 80% of the time, in fury 15% of the time, and in acceptance 5% of the time. Knowing that they congratulated their opponent, now what would you calculate as the probability that the failed party is in denial? And what's the probability they are in fury and in acceptance, respectively?