DSC 40A Fall 2024 - Group Work Session 2

due Monday, October 6th at 11:59PM

Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. **One person** from each group should submit your solutions to Gradescope and **tag all group members** so everyone gets credit.

This worksheet won't be graded on correctness, but rather on good-faith effort. Even if you don't solve any of the problems, you should include some explanation of what you thought about and discussed, so that you can get credit for spending time on the assignment.

In order to receive full credit, you must work in a group of two to four students for at least 50 minutes in your assigned discussion section. You can also self-organize a group and meet outside of discussion section for 80 percent credit. You may not do the groupwork alone.

1 Summation Notation

You can often verify for yourself if something is true about summation notation by "expanding" the summation symbol and seeing if the property holds. For instance, suppose we want to see if it is true that

$$\sum_{i=1}^{n} c \cdot x_i = c \sum_{i=1}^{n} x_i$$

We start by "expanding" $\displaystyle\sum_{i=1}^{n} c \cdot x_i$:

$$\sum_{i=1}^{n} c \cdot x_i = cx_1 + cx_2 + cx_3 + \ldots + cx_n$$

Now we see that the c can be factored out:

$$\begin{aligned} &= c(x_1 + x_2 + x_3 + \ldots + x_n) \\ &= c \sum_{i=1}^{n} x_i. \end{aligned}$$

This is a simple proof that the property is true. On the other hand, we can prove that a property doesn't hold in the same way: by expanding both sides and showing that they are not equal.

Problem 1.

Show that $\displaystyle\sum_{i=1}^{n} (x_i + y_i) = \left(\sum_{i=1}^{n} x_i\right) + \left(\sum_{i=1}^{n} y_i\right)$.

Solution:

$$\begin{aligned}\sum_{i=1}^{n}(x_i + y_i) &= (x_1 + y_1) + (x_2 + y_2) + \cdots + (x_n + y_n) \\ &= (x_1 + x_2 + \ldots x_n) + (y_1 + y_2 + \ldots y_n) \\ &= \left(\sum_{i=1}^{n} x_i\right) + \left(\sum_{i=1}^{n} y_i\right)\end{aligned}$$

Problem 2.

Find a simple expression for $\sum_{i=1}^{n} c$ not involving summation notation. Show that your expression is correct.

Solution: The simple expression is $c * n$, as shown by expanding the sum:

$$\begin{aligned}\sum_{i=1}^{n} c &= c + c + \cdots + c \\ &= c * n.\end{aligned}$$

2 Chaining Inequalities

Suppose we have collected a bunch of numbers, $y_1, \ldots, y_n$. Let's assume, too, that these numbers are in sorted order, so that $y_1 \le y_2 \le \ldots \le y_n$.

The *midpoint* of $y_1, \ldots, y_n$ is the average of the smallest and largest number:

$$\text{midpoint} = \frac{y_1 + y_n}{2}.$$

Intuitively, the midpoint is at most y_n and is at least y_1; it lies somewhere in the middle of these two numbers. We can easily prove this with a *chain* of inequalities.

First, we show that the midpoint is at most y_n. We start with the definition:

$$\text{midpoint} = \frac{y_1 + y_n}{2}$$

We can do anything to the right hand side that makes it bigger, keeping in mind that we're trying to get it to look like y_n. Right now there is y_1 hanging out; can we simply change it to a y_n? Yes! Remember that $y_n \ge y_1$, so this would make the right hand side bigger. Therefore, we have to write $\le$:

$$\le \frac{y_n + y_n}{2}$$

We can simplify this:

$$= \frac{2y_n}{2}$$

Notice that we wrote $=$ on the last line, not $\le$. This is because the line is indeed equal to the one before it.

$$= y_n$$

We have made a chain of inequalities and equalities; this one looks like $=, \le, =, =$. Since $\le$ is the "weakest link" in the chain, the strongest statement we can make is that the midpoint is $\le y_n$, but this is what we wanted to say.

Problem 3.

Prove that the midpoint is $\ge y_1$.

Solution:

$$\begin{aligned} y_1 &\le y_n \\ y_1 + y_1 &\le y_n + y_1 \\ \frac{2y_1}{2} &\le \frac{y_n + y_1}{2} \\ y_1 &\le \mathit{midpoint} \end{aligned}$$

Problem 4.

Suppose $y_1, \ldots, y_n$ are all positive numbers. The *geometric mean* of $y_1, \ldots, y_n$ is defined to be:

$$(y_1 \cdot y_2 \cdots y_n)^{1/n}.$$

Prove that the geometric mean is less than or equal to y_n and greater than or equal to y_1 using a chain of inequalities.

Solution: Assuming the numbers are ordered, let's first show that the geometric mean $\geq y_1$. We know that the below inequalities hold by definition .

$$\begin{aligned} y_1 &\leq y_1 \\ y_1 &\leq y_2 \\ y_1 &\leq y_3 \\ y_1 &\leq y_4 \\ &\dots \\ y_1 &\leq y_n \end{aligned}$$

Since $y_i > 0 \ \forall i$, we can multiply the n inequalities to get

$$y_1 y_1 \dots y_1 \leq y_1 y_2 \dots y_n$$

So,

$$\begin{aligned} y_1^n &\leq y_1 y_2 \dots y_n \\ (y_1^n)^{1/n} &\leq (y_1 y_2 \dots y_n)^{1/n} \\ y_1 &\leq \textit{geometric mean} \end{aligned}$$

You can similarly show that *geometric mean* $\leq y_n$ by using the fact that $y_i \leq y_n$ for $i = 1, 2, \dots n$.

3 Minimizers and Maximizers

We've seen that machine learning problems must first be formulated as mathematical problems. Many of these mathematical problems turn out to be optimization problems: finding the value that minimizes or maximizes a function.

For a function of one variable $f(x)$, a value x^* is said to be a **minimizer** of $f(x)$ if

$$f(x^*) \leq f(x) \quad \text{for all } x.$$

Similarly, x^* is said to be a **maximizer** of $f(x)$ if

$$f(x^*) \geq f(x) \quad \text{for all } x.$$

Notice that a function can have multiple minimizers or maximizers. For example, a constant function like $f(x) = 5$ is minimized at all values of x, and it's also maximized at all values of x!

Problem 5.

Consider the function $g(t) = 2|3t - 4| + 7$. *SUGGESTION: Each of these problems can be proven directly from the definition of minimizer/maximizer above OR using some tool(s) from calculus. You should mix and match your approaches!*

a) Show that $t = 4/3$ is a minimizer of $g(t)$ on the interval $0 \leq t \leq 2$.

> **Solution:** Note that $g(t) = 2|3t-4|+7 \geq 7$ since $|3t-4| \geq 0$ for all t, and thus since $g(4/3) = 7$, we have $g(4/3) \leq g(t)$ for all $t \in [0, 2]$.

b) Find all maximizers of $g(t)$ on the interval $0 \leq t \leq 2$.

> **Solution:** If $0 \leq t \leq 2$, then we can work backwards to bound $g(t)$:
>
> $$\begin{aligned} 0 \leq 3t &\leq 6 \\ -4 \leq 3t - 4 &\leq 2 \\ 0 \leq |3t - 4| &\leq 4 \\ 0 \leq 2|3t - 4| &\leq 8 \\ 7 \leq 2|3t - 4| + 7 &\leq 15 \\ &. \end{aligned}$$
>
> Therefore, $g(t) \leq 15$ for all $t \in [0, 2]$; moreover, $g(t) = 15$ exactly once for $t \in [0, 2]$ when $t = 0$. Therefore $t = 0$ is the only maximizer of g on $[0, 2]$.

c) BONUS: Fix real numbers $a, b \in \mathbb{R}$ such that $a < b$. Find the all maximizers of $g(t)$ on the interval $a \leq t \leq b$.

> **Solution:** You can mimick the last solution or use some calculus. By Fermat's theorem (for extrema), a maximizer t^* must be a critical point for g, where– importantly– its derivative is zero **or undefined**, or a boundary point on the interval. The derivative g' is never zero and is undefined at $t = 4/3$ where, as we saw before, it has a minimizer. Therefore the maximizer(s) are:
>
> $$t^* = \text{argmax}_{t=a,b}(2|3t - 4| + 7),$$
>
> which is a compact way of writing the following: either a, b, or both depending on which one(s)

achieve the maximum of of $2|3t-4|+7$.

Later in the quarter we could use convexity as well.

Problem 6.

Suppose $h(y)$ is a function for a real variable $y \in \mathbb{R}$. Answer each of the following statements with **TRUE** or **FALSE**. If your answer is **FALSE**, write a counter-example. If you answer is **TRUE**, write a proof.

a) If y^* is a minimizer for h, then y^* is a minimizer for h^2.

Solution: FALSE
Counter-example: Let $h(y) = y$ on the interval $[-2, 5]$. Then $y^* = -2$ is a minimizer for h, but y^* is not a minimizer for h^2 since $h(-2)^2 = 4$ and $h(0)^2 = 0 < 4$.

b) If y^* is a minimizer for $|h|$, then y^* is a minimizer for $|h|^2$.

Solution: TRUE
If y^* is a minimizer for $|h|$, the $|h(y^*)| \leq |h(y)|$ for all y in the domain of h. Multiplying both sides by $|h(y^*)| \geq 0$, we have $|h(y^*)|^2 \leq |h(y)||h(y^*)|$ and since $|h(y^*)| \leq |h(y)|$ it follows $|h(y^*)|^2 \leq |h(y)||h(y^*)| \leq |h(y)|^2$ and therefore y^* is also a minimizer for $|h|^2$.

c) If y^* is a maximizer for h, then y^* is a maximizer for $\sqrt{|h|}$.

Solution: FALSE
Counter-example: Let $h(y) = y$ on the interval $[-5, 2]$. Then $y^* = 2$ is a maximizer for h, but y^* is not a maximizer for $\sqrt{|h|}$ since $\sqrt{|h(2)|} = \sqrt{2}$ and $\sqrt{|h(-5)|} = \sqrt{5} > \sqrt{2}$.

d) If y^* is a maximizer for h, then y^* is a maximizer for h^3.

Solution: TRUE
If y^* is a minimizer for h, the $h(y) \leq h(y^*)$ for all y in the domain of h. There are three possible situations we need to consider.

1. If $0 \leq h(y) \leq h(y^*)$, then we can multiply across the inequality by $h(y)^2$ to get $h(y)^3 \leq h(y^*)h(y)^2$. Then, since $0 \leq h(y) \leq h(y^*)$, we can write $h(y^*)h(y)^2 \leq h(y^*)^2 h(y)$ by replacing one of the copies of $h(y)$. One more time, we get $h(y^*)^2 h(y) \leq h(y^*)^3$. Note that you can't do this exact manipulation if either of the terms is negative, since this would involve multiplying an inequality by a negative term which would reverse the direction. Thus, $h(y)^3 \leq h(y^*)^3$.
2. If $h(y) \leq 0 \leq h(y^*)$, then this is easier because $h(y)^3 \leq 0$ and $h(y^*)^3 \geq 0$ (because negative numbers cubed are negative, and positive numbers cubed are positive). Thus, $h(y)^3 \leq h(y^*)^3$.
3. If $h(y) \leq h(y^*) \leq 0$, then $0 \leq -h(y) \leq -h(y^*)$. From part (1) it follows that $0 \leq -h(y)^3 \leq -h(y^*)^3$, but then $h(y)^3 \leq h(y^*)^3 \leq 0$ as desired.

e) If $F : \mathbb{R} \to \mathbb{R}$ is a monotone increasing function (meaning: $F(x) \leq F(y)$ whenever $x \leq y$), and y^* is a minimizer for h, then y^* is a minimizer for $F(h(y))$.

Solution: If y^* is a minimizer for h, then $h(y^*) \leq h(y)$ for all y. Thus since F is monotone, $F(h(y^*)) \leq F(h(y))$ and therefore the claim is TRUE.

f) If $G : \mathbb{R} \to \mathbb{R}$ is a monotone decreasing function (meaning: $G(x) \geq G(y)$ whenever $x \leq y$), and y^* is a maximizer for h, then y^* is a maximizer for $G(h(y))$.

Solution: FALSE - y^* is a MINIMIZER for $G(h(y))$. The proof is very similar to (e).