10-606: Mathematical Foundations for Machine Learning **Fall 2021**

Recitation 3: Worksheet

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

3.1 Linear and Convex Functions

Recall the following definitions

1. f is a *linear function* if f has the properties:
   - $f(u+z) = f(u) + f(z)$
   - $f(cx) = cf(x)$, for a constant c.
2. f is a *convex function* it has the property that

$$f(\alpha u + (1-\alpha)z) \leq \alpha f(u) + (1-\alpha)f(z)$$

Question 1.1 Show that $f(x) = mx$ is a linear function.

Solution. For the first property...

$$f(u+z) = m(u+z) = mu + mz = f(u) + f(z)$$

For the second property...

$$f(cx) = m(cx) = cmx = cf(x)$$

Question 1.2 We saw $f(x) = mx + b$, where $m, b \in \mathbb{R}$, is not linear during lecture. Is it convex? Show why or why not.

Solution. For the left-hand side of $f(\alpha u + (1-\alpha)z) \leq \alpha f(u) + (1-\alpha)f(z)$...

$$f(\alpha u + (1-\alpha)z) = m(\alpha u + (1-\alpha)z) + b = m\alpha u + m(1-\alpha)z + b$$

For the right-hand side...

$$\begin{aligned}\alpha f(u) + (1-\alpha)f(z) &= \alpha(mu + b) + (1-\alpha)(mz + b) \\ &= \alpha mu + \alpha b + (1-\alpha)mz + (1-\alpha)b \\ &= m\alpha u + m(1-\alpha)z + b\end{aligned}$$

Since the left-hand side equals the right-hand side, the inequality holds and we see that f is indeed convex.

Question 1.3 Show that if f is a linear function, it is also a convex function.

Solution.

$$\begin{aligned}f(\alpha u + (1-\alpha)z) &= f(\alpha u) + f((1-\alpha)z) && \text{(Using } f(a+b) = f(a) + f(b) \text{ where } a = \alpha u \text{ and } b = (1-\alpha)z\text{)} \\ &= \alpha f(u) + (1-\alpha)f(z) && \text{(using } f(cx) = cf(x) \text{ two times)}\end{aligned}$$

Since $f(\alpha u + (1-\alpha)z) \leq \alpha f(u) + (1-\alpha)f(z)$ it is convex.

Question 1.4 (Bonus) Let $f(x) = ax^2 + bx + c$ for $a, b, c \in \mathbb{R}$ and $a > 0$. Show that f is convex. *Hint: Note that* $xy \leq x^2 + y^2$.

Solution. For the left-hand side of the inequality...

$$\begin{aligned} f(\alpha u + (1-\alpha)z) &= a(\alpha u + (1-\alpha)z)^2 + b(\alpha u + (1-\alpha)z) + c \\ &= a\alpha^2 u^2 + a(1-\alpha)^2 z^2 + a\alpha(1-\alpha)uz + b\alpha u + b(1-\alpha)z + c \end{aligned}$$

For the right-hand side of the inequality...

$$\begin{aligned} \alpha f(u) + (1-\alpha)f(z) &= \alpha(au^2 + bu + c) + (1-\alpha)(az^2 + bz + c) \\ &= a\alpha u^2 + a(1-\alpha)z^2 + b\alpha u + b(1-\alpha)z + c \end{aligned}$$

Since $b\alpha u + b(1-\alpha)z + c$ is a common term in both of the expression, we just need to show that

$$a\alpha^2 u^2 + a(1-\alpha)^2 z^2 + a\alpha(1-\alpha)uz \leq a\alpha u^2 + a(1-\alpha)z^2$$

or equivalently that

$$\alpha(1-\alpha)uz \leq \alpha u^2 + (1-\alpha)z^2 - \alpha^2 u^2 - (1-\alpha)^2 z^2$$

For the right-hand side of the above equality note that

$$\begin{aligned} &\alpha u^2 + (1-\alpha)z^2 - \alpha^2 u^2 - (1-\alpha)^2 z^2 \\ &= \alpha(u^2 - \alpha u^2) + (1-\alpha)(z^2 - (1-\alpha)z^2) \\ &= \alpha(1-\alpha)u^2 + (1-\alpha)\alpha z^2 \end{aligned}$$

By the hint we see that

$$\alpha(1-\alpha)uz \leq \alpha u^2 + (1-\alpha)z^2 - \alpha^2 u^2 - (1-\alpha)^2 z^2$$

Why is the hint true? Because...

$$\begin{aligned} 0 &\leq (x-y)^2 = x^2 + y^2 - 2xy \\ 2xy &\leq x^2 + y^2 \\ xy &\leq \frac{x^2 + y^2}{2} \leq x^2 + y^2 \end{aligned}$$

3.2 A Refresher on (Scalar) Derivatives

There are a few important derivative rules to remember.

- **The Product Rule.**

$$\frac{d}{dx}\Big(f(x)g(x)\Big) = f'(x)g(x) + f(x)g'(x)$$

- **The Quotient Rule.**

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$$

- **The Chain Rule.**

$$\frac{d}{dx}f(g(x)) = f'(g(x))g'(x)$$

Compute the following derivatives and state which rule (if any) you used for each step:

Question 2.1 Let $f(x) = xe^x$. $f'(x) = \dots$

Remember that $\frac{d}{dx}e^x = e^x$.

Solution.

$$\begin{aligned} f'(x) &= \left(\frac{d}{dx}x\right)e^x + x\left(\frac{d}{dx}e^x\right) \quad \text{(Product Rule)} \\ &= e^x + xe^x \end{aligned}$$

Question 2.2 Let $f(x) = \frac{x}{e^x}$. $f'(x) = \dots$

Solution.

$$\begin{aligned} f'(x) &= \frac{\left(\frac{d}{dx}x\right)e^x - x\left(\frac{d}{dx}e^x\right)}{(e^x)^2} \quad \text{(Quotient Rule)} \\ &= \frac{e^x - xe^x}{e^{2x}} \\ &= \frac{1-x}{e^x} \end{aligned}$$

Question 2.3 Let $f(x) = e^{3x}$. $f'(x) = \dots$

Solution.

$$\begin{aligned} f'(x) &= e^{3x}\left(\frac{d}{dx}3x\right) \quad \text{(Chain Rule)} \\ &= 3e^{3x} \end{aligned}$$

Question 2.4 Let $f(x) = (1+2x)^{563}$. $f'(x) = \dots$

Solution.

$$\begin{aligned} f'(x) &= 563(1+2x)^{562}\left(\frac{d}{dx}(1+2x)\right) \quad \text{(Chain Rule)} \\ &= 1126(1+2x)^{562} \end{aligned}$$

Question 2.5 Let $f(x) = \frac{3x}{(2+3x)^3}$. $f'(x) = \ldots$

Solution.

$$\begin{aligned} f'(x) &= \frac{\left(\frac{d}{dx}3x\right)(2+3x)^3 - 3x\left(\frac{d}{dx}(2+3x)^3\right)}{(2+3x)^6} && \text{(Quotient Rule)} \\ &= \frac{3(2+3x)^3 - 3x\left(3(2+3x)^2\left(\frac{d}{dx}(2+3x)\right)\right)}{(2+3x)^6} && \text{(Chain Rule)} \\ &= \frac{3(2+3x)^3 - 27x(2+3x)^2}{(2+3x)^6} \\ &= \frac{6-18x}{(2+3x)^4} \end{aligned}$$

Question 2.6 Let $f(x) = e^{4xe^{2x}}$. $f'(x) = \ldots$

Solution.

$$\begin{aligned} f'(x) &= e^{4xe^{2x}}\left(\frac{d}{dx}4xe^{2x}\right) && \text{(Chain Rule)} \\ &= e^{4xe^{2x}}\left(\left(\frac{d}{dx}4x\right)e^{2x} + 4x\left(\frac{d}{dx}e^{2x}\right)\right) && \text{(Product Rule)} \\ &= e^{4xe^{2x}}\left(4e^{2x} + 4xe^{2x}\left(\frac{d}{dx}2x\right)\right) && \text{(Chain Rule)} \\ &= e^{4xe^{2x}}\left(4e^{2x} + 8xe^{2x}\right) \end{aligned}$$