ECE/CS/ME 539 – Fall 2024 — Activity 15

1.

Derive the derivatives $\frac{df(x)}{dx}$ of the activation functions $f(x)$ as shown below.

Name of Activation Function $f(x)$	$f(x)$	$f'(x) = \frac{df(x)}{dx}$
ReLU	$\max(0, x)$	$\begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$
Sigmoid	$\frac{1}{1+e^{-x/T}}$	$\frac{1}{T} f(x)(1 - f(x))$
Hyperbolic tangent	$\tanh\left(\frac{x}{T}\right) = \frac{e^{x/T} - e^{-x/T}}{e^{x/T} + e^{-x/T}}$	$\frac{1}{T}\left(1 - f^2(x)\right)$
Inverse tangent	$\frac{2}{\pi}\tan^{-1}\left(\frac{x}{T}\right)$	$\frac{2}{\pi T} \cdot \frac{1}{1+(x/T)^2}$

Answer: Note that the derivatives of activation functions may be computed directly from x or by reusing $f(x)$. Reusing $f(x)$ is useful because it allows us to compute derivatives using the outputs of the activation function, which must be computed anyway, thus saving on compute.

ReLU:

The ReLU activation function is defined as:

$$f(x) = \max(0, x)$$

The derivative of the ReLU function is:

$$\frac{df(x)}{dx} = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$$

Sigmoid:

$$\frac{df(x)}{dx} = \frac{(1/T)\, e^{-x/T}}{\left(1 + e^{-x/T}\right)^2} = f(x) \cdot \frac{1}{T} \cdot \left(1 - \frac{e^{-x/T}}{1 + e^{-x/T}}\right) = \frac{1}{T} f(x)(1 - f(x))$$

Hyperbolic tangent:

$$\frac{d}{dx}\left(\frac{1 - e^{-2x/T}}{1 + e^{-2x/T}}\right) = \frac{(4/T)e^{-2x/T}}{\left(1 + e^{-2x/T}\right)^2} = \frac{1}{T}\left[1 - \left(\frac{1 - e^{-2x/T}}{1 + e^{-2x/T}}\right)^2\right] = \frac{1}{T}\left(1 - f^2(x)\right)$$

Inverse tangent:

$$\frac{d}{dx}\left(\frac{2}{\pi}\tan^{-1}\left(\frac{x}{T}\right)\right) = \frac{2}{\pi} \cdot \frac{d\left(\tan^{-1}\left(\frac{x}{T}\right)\right)}{d\,(x/T)} \cdot \frac{d\,(x/T)}{dx} = \frac{2}{\pi T} \cdot \frac{1}{1 + (x/T)^2}$$

2.

Derive the range (minimum, maximum) of each activation function and its derivative assuming $T = 1$.

Name of Activation Function $f(x)$	$f(x)$	$f'(x) = \frac{df(x)}{dx}$
ReLU	$[0, +\infty)$	$\{0, 1\}$
Sigmoid	$(0, 1)$	$(0, 0.25)$
Hyperbolic tangent	$(-1, +1)$	$(0, 1)$
Inverse tangent	$(-1, +1)$	$(0, \frac{2}{\pi})$

Answer: All four activation functions are non-decreasing functions. Hence,

$$\min f(x) = \lim_{x \to -\infty} f(x) = \begin{cases} 0 & \text{ReLU} \\ 0 & \text{sigmoid} \\ -1 & \text{hyperbolic} \\ -1 & \text{inverse tangent} \end{cases}$$

$$\max f(x) = \lim_{x \to +\infty} f(x) = \begin{cases} +\infty & \text{ReLU} \\ +1 & \text{sigmoid} \\ +1 & \text{hyperbolic} \\ +1 & \text{inverse tangent} \end{cases}$$

ReLU:
The ReLU function ranges from 0 to $+\infty$, and its derivative is either 0 (for $x \leq 0$) or 1 (for $x > 0$).
Sigmoid:

$$0 \leq f(x) \leq 1 \implies f(x) \cdot (1 - f(x)) \geq 0$$

$$\frac{df(x)}{dx} = f(x)(1 - f(x)) = -\big((f(x) - 0.5)^2 - 0.25\big) = 0.25 - (f(x) - 0.5)^2 \leq 0.25$$

Hyperbolic tangent:

$$-1 \leq f(x) \leq +1 \implies 0 \leq f^2(x) \leq 1 \implies 0 \leq 1 - f^2(x) \leq 1$$

Inverse tangent:

$$\frac{2}{\pi} \cdot \frac{1}{1 + x^2} = \frac{2}{\pi}, \quad x = 0$$