A **Probability**

1. **Density function**. Let p be a Gaussian distribution with zero mean and variance of 0.1. Compute the density of p at 0.

$$\text{PDF: } f(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\mu = 0 \quad \sigma^2 = 0.1$$

$$\text{Thus} \quad \sigma = \sqrt{0.1}$$

$$f(0|0,0.1) = \frac{1}{\sqrt{2\pi \cdot 0.1}} e^{\frac{-(0-0)^2}{2\cdot 0.1}}$$

$$= \frac{1}{\sqrt{0.2\pi}}$$

$$\approx 1.2616$$

2. **Conditional probability**. A student is taking a one-hour-time-limit makeup examination. Suppose the probability that the student will finish the exam in less than x hours is x/2, ∀x∈[0,1]. Given that the student is still working after 0.75 hour, what is the conditional probability that the full hour will be used?

assume T = time of the student finish

$$P(T<0.75) = \frac{0.75}{2} = 0.375$$

Thus

$$P(T\geq 0.75) = 1 - P(T<0.75) = 0.625$$

$$P(T=1 \mid T\geq 0.75) = \frac{P(T=1 \cap T\geq 0.75)}{P(T\geq 0.75)}$$

$$= \frac{P(T=1)}{P(T\geq 0.75)}$$

$$= \frac{0.5}{0.625}$$

$$= 0.8$$

3. **Bayes rule.** Consider the probability distribution of you getting sick given the weather in the table below.

Sick?	Weather			
	sunny	rainy	cloudy	snow
yes	0.144	0.02	0.016	0.02
no	0.576	0.08	0.064	0.08

Compute P (sick = yes | Weather = rainy).

Bayes:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

$$P(\text{sick}=\text{yes} \mid \text{Weather}=\text{rainy})$$

$$= \frac{P(\text{sick}=\text{yes})\, P(\text{Weather}=\text{rainy} \mid \text{sick}=\text{yes})}{P(\text{weather}=\text{rainy})}$$

$$= \frac{(0.144+0.02+0.016+0.02)\cdot 0.02}{0.1}$$

$$= 0.04$$

B Calculus and Linear Algebra

1. Compute the derivative of the function f(z) with respect to z(i. e., y/d/s), where

$$f(z) = \frac{1}{1+e^{-z}}$$

$$f(x) = \frac{1}{1+e^{-x}}$$

$$f'(x) = -\frac{(1+e^{-x})'}{(1+e^{-x})^2}$$

$$= -\frac{1}{(1+e^{-x})^2} \cdot (1+e^{-x})'$$

$$= -\frac{1}{(1+e^{-x})^2} \cdot (e^{-x})'$$

$$= -\frac{1}{(1+e^{-x})^2} \cdot (-e^{-x})$$

$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}}$$

$$= \frac{1}{1+e^{-x}} \times \frac{1+e^{-x}-1}{1+e^{-x}}$$

$$= \frac{1}{1+e^{-x}} \times \left(1 - \frac{1}{1+e^{-x}}\right)$$

$$= f(x) \cdot (1 - f(x))$$

2. Compute the derivative of the function f(w) with respect to w_i, where w, x∈RD and

$$f(w) = \frac{1}{1+e^{-w^T x}}$$

$$f(w) = \frac{1}{1+e^{-w^T x}}$$

Given $\sigma(z) = \frac{1}{1+e^{-z}}$

$$\sigma'(z) = \sigma(z)(1-\sigma(z)).$$

$$f(w) = \sigma(w^T x)$$

$$f'(w) = f(w)(1-f(w)) = \frac{df}{dz}$$

$$z = w^T x = \sum_{j=1}^{D} w_j x_j$$

$$\frac{dz}{dw_i} = x_i$$

$$\frac{df}{dw_i} = \frac{df}{dz} \times \frac{dz}{dw_i}$$

$$= f(w)(1-f(w)) \times x_i$$

$$f'(w) = f(w)(1-f(w)) \cdot x_i$$

3. Compute the derivative of the loss function J(w) with respect to w, where

$$J(w) = \frac{1}{2}\sum_{i=1}^{m} |w^T x^{(i)} - y^{(i)}|$$

$$J(w) = \frac{1}{2}\sum_{i=1}^{m} \left|w^T x^{(i)} - y^{(i)}\right|$$

denote . $z^{(i)} = w^T x^{(i)} - y^{(i)}$

subgradient: $\begin{cases} 1 & z>0 \\ -1 & z<0 \\ [-1,1] & z=0 \end{cases}$

Thus . $J(w) = \frac{1}{2}\sum_{i=1}^{m} |z^{(i)}|$

$$J'(w) = \frac{1}{2}\sum_{i=1}^{m} |z^{(i)}|' \cdot z^{(i)\prime}$$

$$= \frac{1}{2}\sum_{i=1}^{m} |z^{(i)}|' \cdot x^{i}$$

$$J'(w) = \begin{cases} \frac{1}{2}\sum_{i=1}^{m} x^{(i)} & \text{if } z^{(i)} > 0 \\ \frac{1}{2}\sum_{i=1}^{m} (-x^{(i)}) & \text{if } z^{(i)} < 0 \\ \left[-\frac{1}{2}\sum_{i=1}^{m} x^{(i)}, \frac{1}{2}\sum_{i=1}^{m} x^{(i)}\right] & \text{if } z^{(i)} = 0 \end{cases}$$

4. Compute the derivative of the loss function J(w) with respect to w, where

$$J(w) = \frac{1}{2}\left[\sum_{i=1}^{m} \left(w^T x^{(i)} - y^{(i)}\right)^2\right] + \lambda ||w||_2^2$$

$$J(w) = \frac{1}{2}\sum_{i=1}^{m} (w^T x - y^i)^2 + \lambda ||w||_2^2$$

Assum $J_1(w) = \frac{1}{2}\sum_{i=1}^{m} (w^T x - y^i)^2$

$$J_2(w) = \lambda ||w||_2^2 = \lambda \sum_{j=1}^{n} w_j^2$$

$$J'(w) = J_1'(w) + J_2'(w)$$

$$= \sum_{i=1}^{m} (w^T x^i - y^i) \cdot x^i + 2\lambda w$$

5. Compute the derivative of the loss function J(w) with respect to w, where

$$J(w) = \Sigma_{i=1}^{m}\left[y^{(i)}\log\left(\frac{1}{1+e^{-w^T x^{(i)}}}\right) + \left(1-y^{(i)}\right)\log\left(1-\frac{1}{1+e^{-w^T x^{(i)}}}\right)\right]$$

Assume $\sigma(z) = \frac{1}{1+e^{-z}}$ (sigmoid)

$\sigma'(z) = \sigma(z)(1-\sigma(z))$

Let $z^{(i)} = w^T x^{(i)}$

$\sigma(z^{(i)}) = \frac{1}{1+e^{-w^T x^{(i)}}} = \frac{1}{1+e^{-z^{(i)}}}$

$$J'(w) = \sum_{i=1}^{m} \left[y^{(i)} \log(\sigma(z^{(i)}))\right]' + \left[(1-y^{(i)}) \log(1-\sigma(z^{(i)}))\right]'$$

$$= \sum_{i=1}^{m} \left[y^{(i)} \cdot \frac{1}{\sigma(z^{(i)})} \cdot \sigma'(z^{(i)}) \cdot x^{i} + (1-y^{(i)}) \cdot \frac{1}{1-\sigma(z^{(i)})} (-\sigma'(z^{(i)})) \cdot x^{i}\right]$$

$$= \sum_{i=1}^{m} \left[y^{(i)} \cdot (1-\sigma(z^{(i)})) \cdot x^{(i)} - (1-y^{(i)}) \cdot \sigma(z^{(i)}) \cdot x^{(i)}\right.$$

$$= \sum_{i=1}^{m} \left[y^{(i)} - y^{(i)}\sigma(z^{(i)}) - \sigma(z^{(i)}) + y^{(i)}\sigma(z^{(i)})\right] x^{(i)}$$

$$= \sum_{i=1}^{m} \left[y^{(i)} - \sigma(z^{(i)})\right] x^{i}$$

6. Compute ▽wf, where f(w)=tanh [wTx].

$$\frac{d}{du} \tanh(u) = 1 - \tanh^2(u)$$

$$\frac{d}{dw} \tanh(w^T x) = (1 - \tanh^2(w^T x)) \cdot \frac{d}{du}(w^T x)$$

$$= (1 - \tanh^2(w^T x)) \cdot x$$

$$\nabla_w f = (1 - \tanh^2(w^T x)) \cdot x$$

7. Find the solution to the system of linear equations given by Ax=b, where

$$A = \begin{pmatrix} 2 & 1 & -1 \\ -3 & -1 & 2 \\ -2 & 1 & 2 \end{pmatrix} \ and\ b = \begin{pmatrix} 8 \\ -11 \\ -3 \end{pmatrix}.$$

$$Ax = b$$

$$x = b \cdot A^{-1}$$

$$A^{-1} = \left(\begin{array}{ccc|ccc} 2 & 1 & -1 & 1 & 0 & 0 \\ -3 & -1 & 2 & 0 & 1 & 0 \\ -2 & 1 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$= \begin{pmatrix} 4 & 3 & -1 \\ -2 & -2 & 1 \\ 5 & 4 & -1 \end{pmatrix}$$

$$x = A^{-1} \cdot B$$

$$= \begin{pmatrix} 4 & 3 & -1 \\ -2 & -2 & 1 \\ 5 & 4 & -1 \end{pmatrix} \cdot \begin{pmatrix} 8 \\ -11 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

$x_1 = 2$ $x_2 = 3$ $x_3 = -1$

C Activation functions (OPTIONAL FOR EXTRA CREDITS)

7. For each of the following activation functions, write their equations and their derivatives. Plot the functions and derivatives, with x∈[-5,5]and y∈[-10,10] plot limits.(No need to submit the code for plots.)

a. Relu

b. Tanh

c. Softmax

d. Sigmoid

e. Leaky ReLU

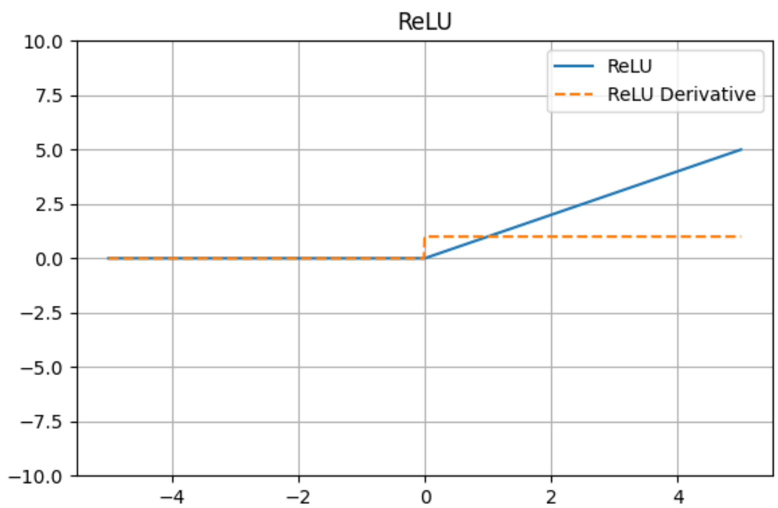
f. Sinc

g. ELU (plot with α=0.3)

ReLU

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

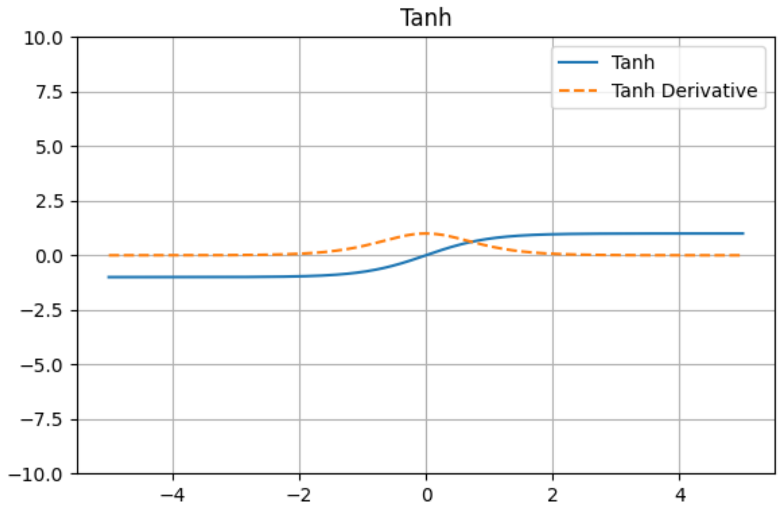
$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$



Tanh

$f(x) = \tanh(x)$

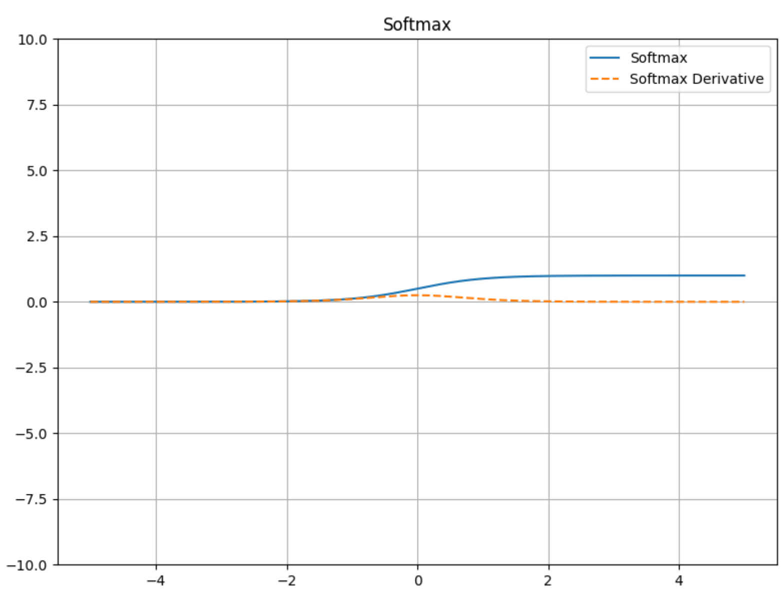
$f'(x) = 1 - \tanh^2(x)$



Softmax

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{m} e^{x_j}}$$

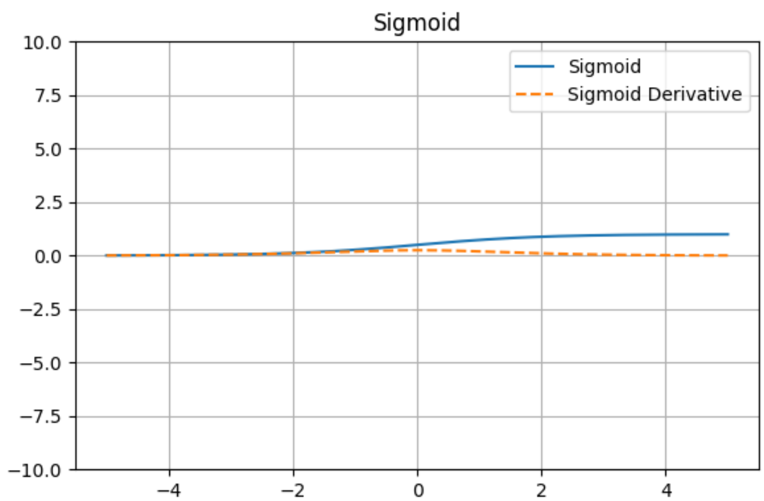
$$f'(x_i) = f(x_i) \times (1 - f(x_i))$$



Sigmoid

$$f(x) = \frac{1}{1+e^{-x}}$$

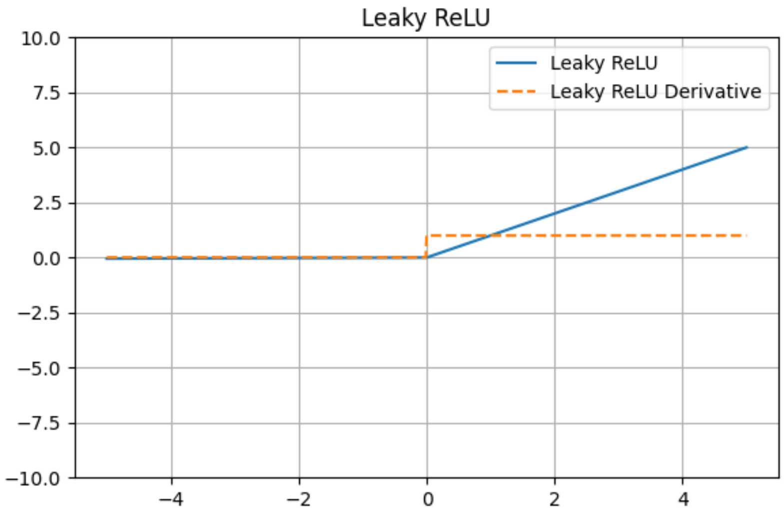
$$f'(x) = f(x) \times (1 - f(x))$$



Leaky ReLU

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{otherwise} \end{cases}$$

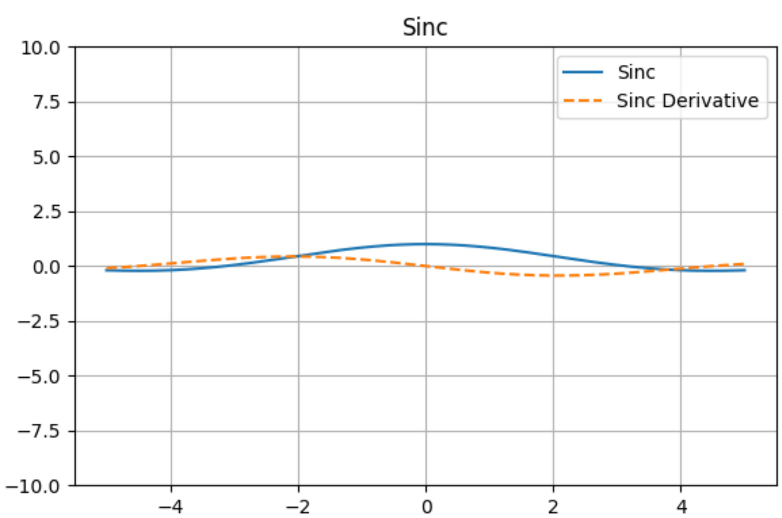
$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ \alpha & \text{otherwise} \end{cases}$$



Sinc

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ \dfrac{\sin(x)}{x} & \text{otherwise} \end{cases}$$

$$f'(x) = \begin{cases} 1 & x = 0 \\ \dfrac{x\cos x - \sin(x)}{x^2} & \text{otherwise} \end{cases}$$



ELU

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha(e^{x}-1) & \text{if } x \leq 0 \end{cases}$$

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ f(x)+\alpha & \text{if } x \leq 0 \end{cases}$$