Problem 1

Prove the following computations:

- For the function $f(z) = tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$ show that $\frac{df}{dz}(z) = 1 - tanh^2(z)$.

Recall:

Quotient Rule

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

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

$$\frac{df}{dz}(z) = \frac{(e^z + e^{-z})(e^z + e^{-z}) - (e^z - e^{-z})(e^z - e^{-z})}{(e^z + e^{-z})^2}$$

$$\frac{df}{dz}(z) = \frac{(e^z + e^{-z})^2 - (e^z - e^{-z})^2}{(e^z + e^{-z})^2}$$

$$\frac{df}{dz}(z) = \frac{(e^z + e^{-z})^2}{(e^z + e^{-z})^2} - \frac{(e^z - e^{-z})^2}{(e^z + e^{-z})^2}$$

$$\frac{df}{dz}(z) = 1 - \frac{(e^z - e^{-z})^2}{(e^z + e^{-z})^2} = 1 - tanh^2(z)$$

- Consider a vector $z = (z_1, \ldots, z_K)$ and the softmax of this vector $a = softmax(z)$ where $a_j = \frac{e^{z_j}}{\sum_{i=1}^K e^{z_i}}$. Find an expression for $\frac{da_j}{dz_j}$ and prove that it is $\frac{\sum_{i=1, j\neq i}^K e^{z_i + z_j}}{(\sum_{i=1}^K e^{z_i})^2}$

$$a_j = \frac{e^{z_j}}{\sum_{i=1}^{K} e^{z_i}}$$

$$\frac{da_j}{dz_j} = \frac{\left(\sum_{i=1}^{K} e^{z_i}\right)\left(e^{z_j}\right) - \left(e^{z_j}\right)\left(e^{z_j}\right)}{\left(\sum_{i=1}^{K} e^{z_i}\right)^2}$$

$$\frac{da_j}{dz_j} = \frac{\sum_{i=1, j\neq i}^{K} \left(e^{z_i}\right)\left(e^{z_j}\right) + \left(e^{z_j}\right)\left(e^{z_j}\right) - \left(e^{z_j}\right)\left(e^{z_j}\right)}{\left(\sum_{i=1}^{K} e^{z_i}\right)^2}$$

$$\frac{da_j}{dz_j} = \frac{\sum_{i=1, j\neq i}^{K} \left(e^{z_i}\right)\left(e^{z_j}\right)}{\left(\sum_{i=1}^{K} e^{z_i}\right)^2}$$

Kate Lassiter

$$\frac{da_j}{dz_j} = \frac{\sum_{i=1, j\neq i}^{K} e^{z_i + z_j}}{\left(\sum_{i=1}^{K} e^{z_i}\right)^2}$$

- Show that for $\sigma(z) = \frac{1}{1+e^{-z}}$ we have $\frac{d\sigma}{dz}(z) = (1-\sigma(z))\sigma(z)$.

$$\sigma(z) = \frac{1}{1+e^{-z}}$$
$$\sigma(z) = \left(1 + e^{-z}\right)^{-1}$$
$$\frac{d\sigma}{dz}(z) = -\left(1 + e^{-z}\right)^{-2}\left(-e^{-z}\right)$$
$$\frac{d\sigma}{dz}(z) = \left(1 + e^{-z}\right)^{-2}\left(e^{-z}\right)$$
$$\frac{d\sigma}{dz}(z) = \frac{\left(e^{-z}\right)}{\left(1+e^{-z}\right)^2}$$
$$\frac{d\sigma}{dz}(z) = \frac{\left(e^{-z}\right)}{\left(1+e^{-z}\right)\left(1+e^{-z}\right)}$$
$$\frac{d\sigma}{dz}(z) = \frac{(1)\ \left(e^{-z}\right)}{\left(1+e^{-z}\right)\left(1+e^{-z}\right)}$$
$$\frac{d\sigma}{dz}(z) = \frac{(1)\ \left(1+e^{-z}-1\right)}{\left(1+e^{-z}\right)\left(1+e^{-z}\right)}$$
$$\frac{d\sigma}{dz}(z) = \frac{(1)}{\left(1+e^{-z}\right)} \times \frac{\left(1+e^{-z}\right)}{\left(1+e^{-z}\right)} - \frac{1}{\left(1+e^{-z}\right)}$$
$$\frac{d\sigma}{dz}(z) = \frac{1}{\left(1+e^{-z}\right)} \times \left(1 - \frac{1}{\left(1+e^{-z}\right)}\right)$$
$$\frac{d\sigma}{dz}(z) = \sigma(z) \times (1-\sigma(z))$$

- Show that $tanh(z) = 2\sigma(2z) - 1$.

$$tanh(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$$
$$tanh(z) = \frac{e^{z} - 2e^{-z} + e^{-z}}{e^{z} + e^{-z}}$$
$$tanh(z) = \frac{-2e^{-z} + e^{z} + e^{-z}}{e^{z} + e^{-z}}$$
$$tanh(z) = \frac{-2e^{-z}}{e^{z} + e^{-z}} + \frac{e^{z} + e^{-z}}{e^{z} + e^{-z}}$$
$$tanh(z) = \frac{-2e^{-z}}{e^{z} + e^{-z}} + 1$$
$$tanh(z) = \frac{-2}{e^{2z} + 1} + 1$$

Recall:

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

$$tanh(z) = -2\left(\frac{1}{e^{-z(-2)}+1}\right) + 1$$
$$tanh(z) = -2\sigma(-2z) + 1$$
$$tanh(z) = -2(\sigma(-z)) + 1$$

<u>Recall:</u>

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

$$\sigma(x) - 1 = -\sigma(-x)$$

$$tanh(z) = 2(-\sigma(-z)) + 1$$
$$tanh(z) = 2(\sigma(z) - 1) + 1$$
$$tanh(z) = 2\sigma(2z) - 1$$

Problem 2

Review the XOR example from class. In the XOR example with a neural network, we picked γ and ν to be specific values. Suppose we use σ and not $ReLU$. Can you find γ and ν that work? Prove this. Do this by smart guess and check. You can write a small Python program to get you the values you need.

Kate Lassiter

Problem 2

Assuming: $B^{[1]} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $\alpha^{[1]} = (0, -1)$

$$a^{[1]} = \text{sigmoid}(z^{[1]})$$

$$a^{[1]} = \text{sigmoid}(B^{[1]T}x + \alpha^{[1]})$$

$$x = (1,0) \to a^{[1]} = \text{sigmoid}(B_1^{[1]} + B_2^{[1]}, B_1^{[1]} + B_2^{[2]} - 1)$$

$$= \text{sigmoid}(1+0, 1+0-1)$$

$$= \text{sigmoid}(1, 0) \to = \boxed{\left(\frac{1}{1+e^{-1}}, \frac{1}{2}\right)}$$

$$x = (0,0) \to a^{[1]} = \text{sigmoid}(0+0, 0+0-1)$$

$$= \text{sigmoid}(0, -1) \to = \boxed{\left(\frac{1}{2}, \frac{1}{1+e}\right)}$$

$$x = (0,1) \to a^{[1]} = \text{sigmoid}(0+1, 0+1-1)$$

$$= \text{sigmoid}(1, 0) \to \boxed{\left(\frac{1}{1+e^{-1}}, \frac{1}{2}\right)}$$

$$x = (1,1) \to a^{[1]} = \text{sigmoid}(1+1, 1+1-1)$$

$$= \text{sigmoid}(2, 1) \to \boxed{\left(\frac{1}{1+e^{-2}}, \frac{1}{1+e^{-1}}\right)}$$

$$a^{[2]} = \text{sigmoid}(\underbrace{B^{[2]T}a^{[1]} + \alpha^{[2]}}_{z^{[2]}})$$

$$x = (1,0) = \text{sigmoid}\left(\frac{1}{1+e^{-1}} + \frac{1}{2} + \alpha_1^{[2]}, \frac{1}{1+e^{-1}} + \frac{1}{2} + \alpha_2^{[2]}\right)$$

$$x = (0,0) = \text{sigmoid}\left(\frac{1}{2} + \frac{1}{1+e} + \alpha_1^{[2]}, \frac{1}{2} + \frac{1}{1+e} + \alpha_1^{[2]}\right)$$

$$x = (0,1) = \text{sigmoid}\left(\frac{1}{1+e^{-1}} + \frac{1}{2} + \alpha_1^{[2]}, \frac{1}{1+e^{-1}} + \frac{1}{2} + \alpha_1^{[2]}\right)$$

$$x = (1,1) = \text{sigmoid}\left(\frac{1}{1+e^{-2}} + \frac{1}{1+e^{-1}} + \alpha_1^{[2]}, \frac{1}{1+e^{-2}} + \frac{1}{1+e^{-1}} + \alpha_1^{[2]}\right)$$

$$a^{[2]} = \text{sigmoid}(z^{[2]})$$

$$p(y=2|x) = \frac{1}{1+e^{-(B_2^{[2]} - B_1^{[2]})^T a^{[1]} + \alpha_2^{[2]} - \alpha_1^{[2]})}} > \frac{1}{2}$$

$$p(y=1|x) = 1 - \frac{1}{1+e^{-(B_2^{[2]} - B_1^{[2]})^T a^{[1]} + \alpha_2^{[2]} - \alpha_1^{[2]}}} > \frac{1}{2}$$

set $\to \gamma = B_2^{[2]} - B_1^{[2]} \in \mathbb{R}^2$ and $v = \alpha_2^{[2]} - \alpha_1^{[2]} \in \mathbb{R}$

$$p(y=2|x) = \frac{1}{1+e^{-(\gamma^T a^{[1]} + v)}}$$

$$y = 2, \gamma^T a^{[1]} + v > 0$$

$$y = 1, \gamma^T a^{[1]} + v < 0$$

must meet these conditions:

$$\frac{1}{2} = \gamma^T\left(\frac{1}{1+e^{-1}}, \frac{1}{2}\right) + v > 0 \text{ for } y=2, x=(1,0)$$

$$-\frac{1}{2} = \gamma^T\left(\frac{1}{2}, \frac{1}{1+e}\right) + v < 0 \text{ for } y=1, x=(0,0)$$

$$\frac{1}{2} = \gamma^T\left(\frac{1}{1+e^{-1}}, \frac{1}{2}\right) + v \geq 0 \text{ for } y=2, x=(0,1)$$

$$-\frac{1}{2} = \gamma^T\left(\frac{1}{1+e^{-2}}, \frac{1}{1+e^{-1}}\right) + v < 0 \text{ for } y=1, x=(1,0)$$

{tried getting values with python notebook, could not return a valid γ and v.)

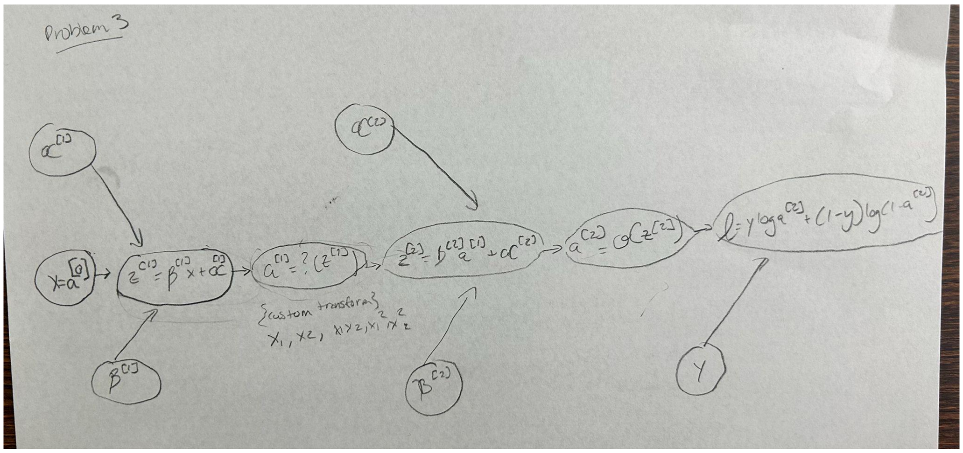
This found no valid results

```
[2] import numpy as np
    as_=np.arange(-10, 10, 0.25)
    bs_=np.arange(-10, 10, 0.25)
    cs_=np.arange(-10, 10, 0.25)

[19] g=list([(1/(1 + math.exp(-1))),(1/(1 + math.exp(-2))),(1/(1 + math.exp(1))),(1/(1 + math.exp(2))),(1 + math.exp(-1)),(1 + math.exp(-2)),(1 + math.exp(1)),

[22] as_=np.append( as_,g)
     bs_=np.append( bs_,g)
     cs_=np.append( cs_,g)

for a in as_:
    for b in bs_:
        for c in cs_:
            success=0
            if round(a*(1/(1 + math.exp(-1)))+b*0.5+c,1)==0.5:
             # print('pass1')
              success=success +1
            if round( a*0.5 +b*(1/(1 + math.exp(1)))+c  ,1) ==-0.5:
              # print('pass2')
               success=success +1
            if round( a*(1/(1 + math.exp(-2))) +b*(1/(1 + math.exp( -1)))+c,1)==-0.5:
               #print('pass3')
               success=success +1
            if success==3:
              print(a,b,c)
```

Kate Lassiter

Problem 3

Suppose we have a neural network as in class and the output of the layer $a^{[1]}$ is $(x_1, x_2, x_1x_2, x_1^2, x_2^2)$ where $x = (x_1, x_2)$ is the input. Recall that for XOR we have (x_1, x_2) maps to y via $y = x_1 + x_2 - 2x_1x_2$ so that $(0, 0)$ maps to 0 and $(1, 0)$ maps to 1 (see Lecture). Consider $z^{[2]} = \beta^{[2]}a^{[1]} + \alpha^{[2]}$ and $a^{[2]} = \sigma(z^{[2]})$ and how we want $a^{[2]} > 1/2$ if $y = 1$ and $1 - a^{[2]} > 1/2$ if $y = 0$. Can you specify $\beta^{[2]}$ and $\alpha^{[2]}$ that make this happen? Notice $\beta^{[2]} \in \mathbb{R}^5$ and $\alpha^{[2]} \in \mathbb{R}$. Do this by smart guess and check. You can write a small Python program to get you the values you need.



Inequalities

$\sigma(\beta_1+\beta_4+a) > \frac{1}{2} = \frac{1}{1+e^{-(\beta_1+\beta_4+a)}} \geq \frac{1}{2}$ ← $y=1,\ x=(1,0) \rightarrow a^{[1]} = (1,0,0,1,0) = \beta_1+\beta_4+a$

$\sigma(\beta_2+\beta_5+a) > \frac{1}{2} = \frac{1}{1+e^{-(\beta_2+\beta_5+a)}} \geq \frac{1}{2}$ ← $y=1,\ x=(0,1) \rightarrow a^{[1]} = (0,1,0,0,1) = \beta_2+\beta_5+a$

$\sigma(a) < \frac{1}{2} = \frac{1}{1+e^{-(a)}} < \frac{1}{2}$ ← $y=0,\ x=(0,0) \rightarrow a^{[1]} = (0,0,0,0,0) = a$

$\sigma(\beta_1+\beta_2+\dots\beta_5) < \frac{1}{2} = \frac{1}{1+e^{-(\beta_1+\dots+\beta_5)}} < \frac{1}{2}$ ← $y=0,\ x=(1,1) \rightarrow a^{[1]} = (1,1,1,1,1) = \beta_1+\beta_2+\beta_3+\beta_4+\beta_5+\alpha$

Transform | **Parameters**

$z^{[2]} = \beta^{[2]}a^{[1]} + \alpha^{[2]}$

$\sigma(z^{[2]}) = \frac{1}{1+e^{-z^{[2]}}} = \frac{1}{1+e^{-(\beta^{[2]}a^{[1]}+\alpha^{[2]})}}$

Sigmoid

Inequalities to Satisfy:

$\beta_1+\beta_4+a \geq \frac{1}{2}$

$\beta_2+\beta_5+a \geq \frac{1}{2}$

$a < \frac{1}{2}$

$\beta_1+\beta_2+\beta_3+\beta_4+\beta_5+a < \frac{1}{2}$

Because β_3 is a free parameter, there are infinitely many parameters that could make this work.

Example: $\alpha = -\frac{1}{2} < \frac{1}{2}$ ✓, $\beta_1 = 1$ and $\beta_4 = 2 \rightarrow \beta_1+\beta_4+\alpha \geq \frac{1}{2} \rightarrow 1+2-\frac{1}{2} \geq \frac{1}{2}$ ✓

$\beta_2 = 1$ and $\beta_5 = 2 \rightarrow \beta_2+\beta_5+a \geq \frac{1}{2} \rightarrow 1+2-\frac{1}{2} \geq \frac{1}{2}$ ✓

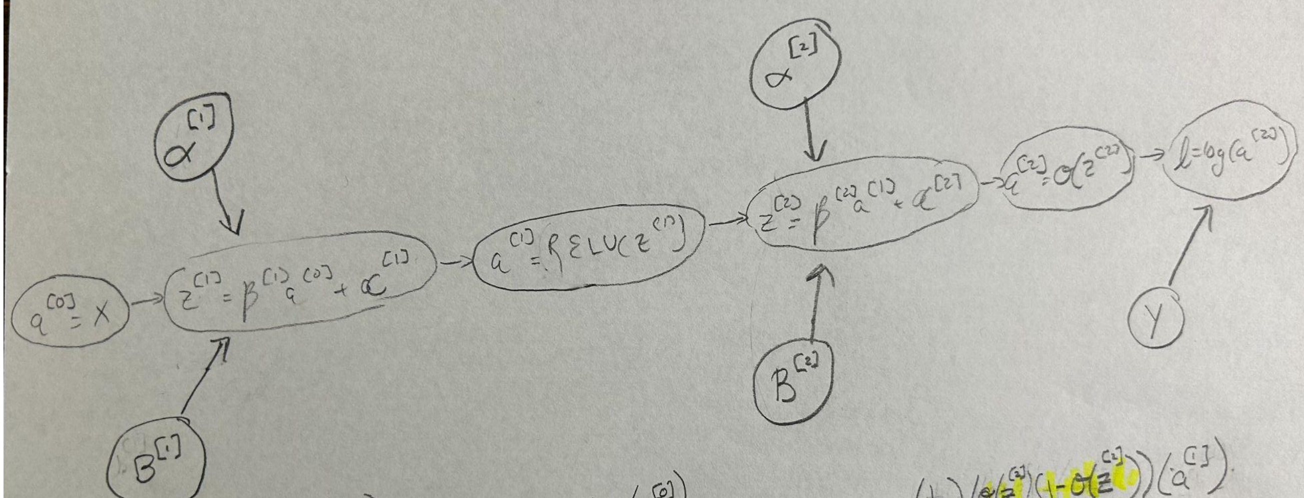
$\beta_3 = -6 \rightarrow \beta_1+\beta_2+\beta_3+\beta_4+\beta_5+a < \frac{1}{2} \rightarrow 1+1-6+2+2-\frac{1}{2} < \frac{1}{2}$ ✓

Kate Lassiter

Problem 4

Suppose we use a *ReLU* so that the recursions are $a^{[0]} = x$, $z^{[1]} = \beta^{[1]}a^{[0]} + \alpha^{[1]}$, $a^{[1]} = ReLU(z^{[1]})$, $z^{[2]} = \beta^{[2]}a^{[1]} + \alpha^{[2]}$ and then finally $a^{[2]} = \sigma(z^{[2]})$ and $\ell = \log(a^{[2]})$ (i.e. we assume $y = 1$). What are the derivatives of ℓ with respect to $\beta^{[1],[2]}$ and $\alpha^{[1],[2]}$. For each variable, when will they be zero? Give some sufficient conditions in terms of the $z^{[1]}$ or $z^{[2]}$ variables.

Problem 4



① $$\frac{d\ell}{dB^{[1]}} = \underbrace{\frac{d\ell}{da^{[2]}}}_{\left(\frac{1}{a^{[2]}}\right)} \frac{da^{[2]}}{dz^{[2]}} \frac{dz^{[2]}}{da^{[1]}} \frac{da^{[1]}}{dz^{[1]}} \underbrace{\frac{dz^{[1]}}{dB^{[1]}}}_{(a^{[0]})}$$

② $$\frac{d\ell}{d\alpha^{[1]}} = \frac{d\ell}{da^{[2]}} \frac{da^{[2]}}{dz^{[2]}} \frac{dz^{[2]}}{da^{[1]}} \frac{da^{[1]}}{dz^{[1]}} \frac{dz^{[1]}}{d\alpha^{[1]}}$$

$$\left(\frac{1}{a^{[2]}}\right)\underbrace{(\sigma(z^{[2]}))(1-\sigma(z^{[2]}))}_{\substack{\lim_{z\to\infty}\sigma(z)=1 \\ \lim_{z\to-\infty}\sigma(z)=0}}(B^{[2]})\underbrace{I_{(0,\infty)}}_{\text{if } z<0,\ \text{this is } 0}(z^{[1]})(1)$$

Conditions:

$$\frac{d\sigma}{dz} = \sigma(z)(1-\sigma(z)) = 0(1-0) = 0 \checkmark = 1(1-1) = 0 \checkmark$$

③ $$\frac{d\ell}{dB^{[2]}} = \frac{d\ell}{da^{[2]}} \frac{da^{[2]}}{dz^{[2]}} \frac{dz^{[2]}}{dB^{[2]}}$$

$$\left(\frac{1}{a^{[2]}}\right)(\sigma(z^{[2]})(1-\sigma(z^{[2]})))(a^{[1]})$$

④ $$\frac{d\ell}{d\alpha^{[2]}} = \frac{d\ell}{da^{[2]}} \frac{da^{[2]}}{dz^{[2]}} \frac{dz^{[2]}}{d\alpha^{[2]}}$$

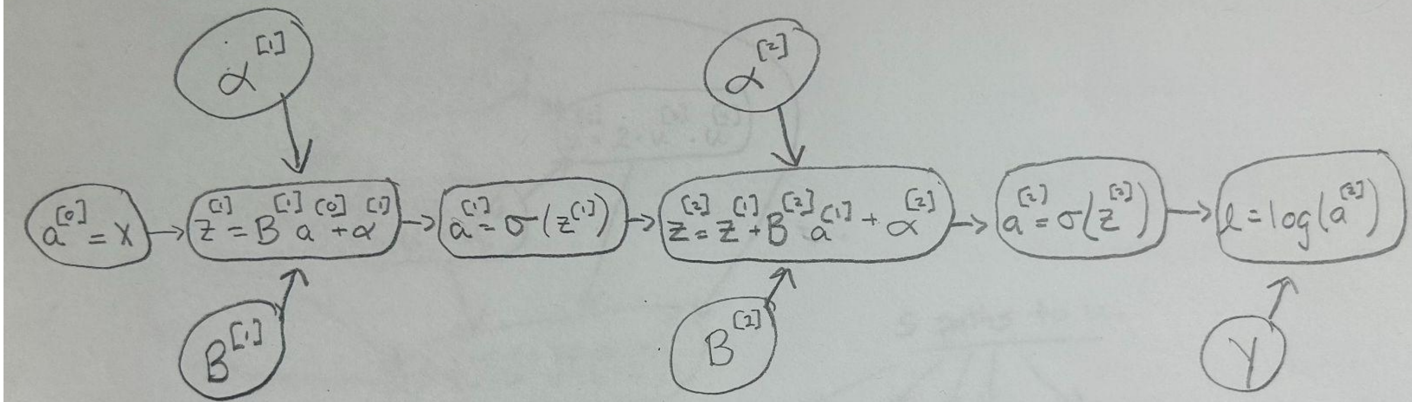
$$\left(\frac{1}{a^{[2]}}\right)\underbrace{(\sigma(z^{[2]}))(1-\sigma(z^{[2]}))}_{\substack{\lim_{z\to\infty}(z)=1 \\ \lim_{z\to-\infty}\sigma(z)=0}}(1)$$

$$\frac{d\sigma}{dz} = \sigma(z)(1-\sigma(z)) = 0(1-0) = 0 \checkmark = 1(1-1) = 0 \checkmark$$

Problem 5

Suppose we have $a^{[0]} = x$, $z^{[1]} = \beta^{[1]}a^{[0]} + \alpha^{[1]}$, $a^{[1]} = \sigma(z^{[1]})$, $z^{[2]} = z^{[1]} + \beta^{[2]}a^{[1]} + \alpha^{[2]}$ and then finally $a^{[2]} = \sigma(z^{[2]})$ and again $\ell = \log(a^{[2]})$. What are the derivatives of ℓ with respect to $\beta^{[1],[2]}$ and $\alpha^{[1],[2]}$. For each variable, when will they be zero? Give some sufficient conditions in terms of the $z^{[1]}$ or $z^{[2]}$ variables. Also, draw the computational graph.

Problem 5



① $\dfrac{d\ell}{dB^{[1]}} = \underbrace{\dfrac{d\ell}{da^{[2]}}}_{\left(\frac{1}{a^{[2]}}\right)} \dfrac{da^{[2]}}{dz^{[2]}} \dfrac{dz^{[2]}}{da^{[1]}} \dfrac{da^{[1]}}{dz^{[1]}} \underbrace{\dfrac{dz^{[1]}}{dB^{[1]}}}_{(a^{[0]})}$

③ $\dfrac{d\ell}{dB^{[2]}} = \dfrac{d\ell}{da^{[2]}} \dfrac{da^{[2]}}{dz^{[2]}} \dfrac{dz^{[2]}}{dB^{[2]}}$
$\left(\frac{1}{a^{[2]}}\right)\left(\sigma(z^{[2]})(1-\sigma(z^{[2]}))\right)(a^{[1]})$

② $\dfrac{d\ell}{d\alpha^{[1]}} = \dfrac{d\ell}{da^{[2]}} \dfrac{da^{[2]}}{dz^{[2]}} \dfrac{dz^{[2]}}{da^{[1]}} \dfrac{da^{[1]}}{dz^{[1]}} \dfrac{dz^{[1]}}{d\alpha^{[1]}}$
$\left(\frac{1}{a^{[2]}}\right)\left(\sigma(z^{[2]})(1-\sigma(z^{[2]}))\right)(B^{[2]})\left(\sigma(z^{[1]})(1-\sigma(z^{[1]}))\right)(1)$

④ $\dfrac{d\ell}{d\alpha^{[2]}} = \dfrac{d\ell}{da^{[2]}} \dfrac{da^{[2]}}{dz^{[2]}} \dfrac{dz^{[2]}}{d\alpha^{[2]}}$
$\left(\frac{1}{a^{[2]}}\right)\left(\sigma(z^{[2]})(1-\sigma(z^{[2]}))\right)(1)$

zero conditions:

$\lim_{z\to\infty}\sigma(z) = 1$ $\quad \lim_{z\to-\infty}\sigma(z) = 0$

$\dfrac{d\sigma}{dz} = \sigma(z)(1-\sigma(z))$
$= 0(1-0) = 0$ ✓
$= 1(1-1) = 0$ ✓

$\lim_{z\to\infty}\sigma(z) = 1$ $\quad \lim_{z\to-\infty}\sigma(z) = 0$

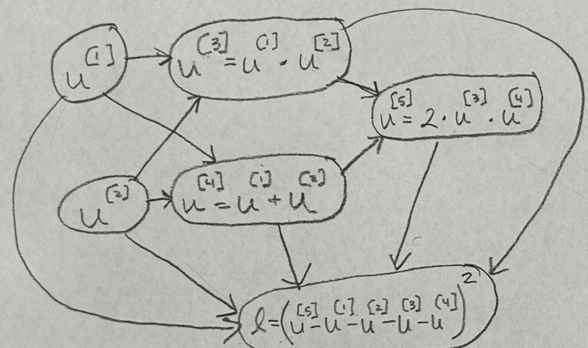
$\dfrac{d\sigma}{dz} = \sigma(z)(1-\sigma(z))$
$= 0(1-0) = 0$ ✓
$= 1(1-1) = 0$ ✓ $\quad \lim_{z\to\infty}\sigma(z) = 1$ $\quad \lim_{z\to-\infty}\sigma(z) = 0$

Kate Lassiter

Problem 6

Suppose we have equations $u^{[3]} = u^{[1]} \times u^{[2]}$, $u^{[4]} = u^{[1]} + u^{[2]}$ and $u^{[5]} = 2 \times u^{[3]} \times u^{[4]}$. Suppose $\mathcal{L} = (u^{[5]} - u^{[1]} - u^{[2]} - u^{[3]} - u^{[4]})^2$ is what we'd like to minimize. Draw the computational graph. Suppose $u^{[1]} = 3$ and $u^{[2]} = 4$, find the partials $\{\frac{\partial \mathcal{L}}{\partial u^{[i]}}\}_{i=1}^{2}$. Write expressions for these partials in terms of local partials. To get the numerical answer: if you want, you can submit PyTorch code with this which sets up the above as tensors and then gets the gradients.

Problem 6



5 paths to u_1

$$\frac{dl}{du^{[1]}} = \left(\frac{dl}{du^{[1]}}\right) + \left(\frac{dl}{du^{[4]}}\frac{du^{[4]}}{du^{[1]}}\right) + \left(\frac{dl}{du^{[5]}}\frac{du^{[5]}}{du^{[3]}}\frac{du^{[3]}}{du^{[1]}}\right) + \left(\frac{dl}{du^{[5]}}\frac{du^{[5]}}{du^{[4]}}\frac{du^{[4]}}{du^{[1]}}\right) + \left(\frac{dl}{du^{[3]}}\frac{du^{[3]}}{du^{[1]}}\right)$$

$$= -2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]}) + -2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]}) + 2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]})(2u^{[4]})(u^{[2]}) + 2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]})(2u^{[3]}) + -2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]})(u^{[2]})$$

$$\frac{dl}{u^{[1]}}(3,4) = -2(168-3-4-12-7) - 2(168-3-4-12-7) + 2(168-3-4-12-7)(2\cdot 7)(4) + 2(168-3-4-12-7)(2\cdot 12) - 2(168-3-4-12-7)(4)$$

$$= \boxed{21{,}016}$$

5 paths to u_2

$$\frac{dl}{du^{[2]}} = \left(\frac{dl}{du^{[2]}}\right) + \left(\frac{dl}{du^{[4]}}\frac{du^{[4]}}{du^{[2]}}\right) + \left(\frac{dl}{du^{[5]}}\frac{du^{[5]}}{du^{[3]}}\frac{du^{[3]}}{du^{[2]}}\right) + \left(\frac{dl}{du^{[5]}}\frac{du^{[5]}}{du^{[4]}}\frac{du^{[4]}}{du^{[2]}}\right) + \left(\frac{dl}{du^{[3]}}\frac{du^{[3]}}{du^{[2]}}\right)$$

$$= -2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]}) + -2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]}) + 2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]})(2u^{[4]})(u^{[1]}) + 2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]})(2u^{[3]}) + -2(u^{[5]}-u^{[1]}-u^{[2]}-u^{[3]}-u^{[4]})(u^{[1]})$$

$$\frac{dl}{u^{[2]}}(3,4) = \boxed{17{,}324}$$

Kate Lassiter

In [ ]:
```
import torch
```

In [11]:
```
u1 = torch.tensor([[ 3.0]], requires_grad=True)
u2 = torch.tensor([[ 4.0]], requires_grad=True)

u3 = u1* u2
u4 = u1 + u2
u5 = 2* (u3* u4)

loss = (u5-u1 -u2 -u3 -u4)**2
loss.backward()

print("loss: ", loss)
print("u1.grad: ", u1.grad)
print("u2.grad: ", u2.grad)
```

```
loss:  tensor([[20164.]], grad_fn=<PowBackward0>)
u1.grad:  tensor([[21016.]])
u2.grad:  tensor([[17324.]])
```

In [13]:
```
# manually calculated gradients:
# there are five paths each to u1, u2.
#d/du1=-2*(u5-u1-u2-u3-u4) - 2*(u5-u1-u2-u3-u4) + (2*(u5-u1-u2-u3-u4)*2*u4*u2) + (2*(u5-u1-u2-u3-u4)*2*u3) + (-2*(u5
-2*(168 -3-4-12-7)-2*(168 -3-4-12-7) + (2*(168 -3-4-12-7)*2*7*4) + (2*(168 -3-4-12-7)*2*12) + (-2*(168 -3-4-12-7)*4)
```

Out[13]: 21016

In [14]:
```
#d/du2=-2*(u5-u1-u2-u3-u4) - 2*(u5-u1-u2-u3-u4) + (2*(u5-u1-u2-u3-u4)*2*u4*u1) + (2*(u5-u1-u2-u3-u4)*2*u3) + (-2*(u5
-2*(168 -3-4-12-7)-2*(168 -3-4-12-7) + (2*(168 -3-4-12-7)*2*7*3) + (2*(168 -3-4-12-7)*2*12) + (-2*(168 -3-4-12-7)*3)
```

Out[14]: 17324