CENG 499 - Introduction to Machine Learning
Fall 2022
Homework 1

Necat Kılıçarslan

November 2022

1 PART 1

First lets solve the regression problem.
In order to solve this problem in a clean manner, firstly lets calculate the sigmoid functions of the hidden layers step by step.

$$\frac{d\sigma(x)}{dx} = \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}} = \frac{1}{1+e^{-x}} \cdot \left(1 - \frac{1}{1+e^{-x}}\right) = \sigma(x) \cdot (1-\sigma(x))$$

$$O_1^{(1)} = \sigma(O_0^{(0)} \cdot a_{01}^{(0)} + O_1^{(0)} \cdot a_{11}^{(0)} + O_2^{(0)} \cdot a_{21}^{(0)})$$

$$O_2^{(1)} = \sigma(O_0^{(0)} \cdot a_{02}^{(0)} + O_1^{(0)} \cdot a_{12}^{(0)} + O_2^{(0)} \cdot a_{22}^{(0)})$$

$$O_3^{(1)} = \sigma(O_0^{(0)} \cdot a_{03}^{(0)} + O_1^{(0)} \cdot a_{13}^{(0)} + O_2^{(0)} \cdot a_{23}^{(0)})$$

Since in the question state that $O_0^0 = 1, O_1^0 = x1, O_0^2 = x2$, we can rewrite these 3 equations in the following manner:

$$O_1^{(1)} = \sigma(O_0^{(0)} \cdot a_{01}^{(0)} + O_1^{(0)} \cdot a_{11}^{(0)} + O_2^{(0)} \cdot a_{21}^{(0)}) = \sigma(1 \cdot a_{01}^{(0)} + x1 \cdot a_{11}^{(0)} + x2 \cdot a_{21}^{(0)})$$

$$O_2^{(1)} = \sigma(O_0^{(0)} \cdot a_{02}^{(0)} + O_1^{(0)} \cdot a_{12}^{(0)} + O_2^{(0)} \cdot a_{22}^{(0)}) = \sigma(1 \cdot a_{02}^{(0)} + x1 \cdot a_{12}^{(0)} + x2 \cdot a_{22}^{(0)})$$

$$O_3^{(1)} = \sigma(O_0^{(0)} \cdot a_{03}^{(0)} + O_1^{(0)} \cdot a_{13}^{(0)} + O_2^{(0)} \cdot a_{23}^{(0)}) = \sigma(1 \cdot a_{03}^{(0)} + x1 \cdot a_{13}^{(0)} + x2 \cdot a_{23}^{(0)})$$

Now lets calculate the derivative of sigmoid function.

$$\frac{d\sigma(O_1^{(1)})}{d(O_1^1)} = \frac{1}{1+e^{-(1 \cdot a_{01}^{(0)} + x1 \cdot a_{11}^{(0)} + x2 \cdot a_{21}^{(0)})}} \cdot \left(1 - \frac{1}{1+e^{-(1 \cdot a_{01}^{(0)} + x1 \cdot a_{11}^{(0)} + x2 \cdot a_{21}^{(0)})}}\right)$$

$$\frac{d\sigma(O_2^{(1)})}{d(O_2^1)} = \frac{1}{1+e^{-(1 \cdot a_{02}^{(0)} + x1 \cdot a_{12}^{(0)} + x2 \cdot a_{22}^{(0)})}} \cdot \left(1 - \frac{1}{1+e^{-(1 \cdot a_{02}^{(0)} + x1 \cdot a_{12}^{(0)} + x2 \cdot a_{22}^{(0)})}}\right)$$

$$\frac{d\sigma(O_3^{(1)})}{d(O_3^1)} = \frac{1}{1+e^{-(1 \cdot a_{03}^{(0)} + x1 \cdot a_{13}^{(0)} + x2 \cdot a_{23}^{(0)})}} \cdot \left(1 - \frac{1}{1+e^{-(1 \cdot a_{03}^{(0)} + x1 \cdot a_{13}^{(0)} + x2 \cdot a_{23}^{(0)})}}\right)$$

$$O_0^{(2)} = a_{00}^{(1)} + a_{10}^{(1)} O_1^{(1)} + a_{20}^{(1)} O_2^{(1)} + a_{30}^{(1)} O_3^{(1)}$$

$$= a_{00}^{(1)} + a_{10}^{(1)} \cdot \left(\frac{1}{1+e^{-(a_{01}^{(0)} + a_{11}^{(0)} x1 + a_{21}^{(0)} x2)}}\right) + a_{20}^{(1)} \cdot \left(\frac{1}{1+e^{-(a_{02}^{(0)} + a_{12}^{(0)} x1 + a_{22}^{(0)} x2)}}\right) + a_{30}^{(1)} \cdot \left(\frac{1}{1+e^{-(a_{03}^{(0)} + a_{13}^{(0)} x1 + a_{23}^{(0)} x2)}}\right)$$

Now we can calculate all the weights much more easily.

- Lets calculate $a_{00}^{(1)}$

$$a_{00}^{(1)\prime} = a_{00}^{(1)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{00}^{(1)}}$$

$$a_{00}^{(1)\prime} = a_{00}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \frac{\partial (O_0^{(2)})}{\partial a_{00}^{(1)}}$$

$$a_{00}^{(1)\prime} = a_{00}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)})$$

- Lets calculate $a_{10}^{(1)}$

$$a_{10}^{(1)\prime} = a_{10}^{(1)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{10}^{(1)}}$$

$$a_{10}^{(1)\prime} = a_{10}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \frac{\partial (O_0^{(2)})}{\partial a_{10}^{(1)}}$$

$$a_{10}^{(1)\prime} = a_{10}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)} x1 + a_{21}^{(0)} x2)}} \right)$$

- Lets calculate $a_{20}^{(1)}$

$$a_{20}^{(1)\prime} = a_{20}^{(1)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{20}^{(1)}}$$

$$a_{20}^{(1)\prime} = a_{20}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \frac{\partial (O_0^{2})}{\partial a_{20}^{(1)}}$$

$$a_{20}^{(1)\prime} = a_{20}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)} x1 + a_{22}^{(0)} x2)}} \right)$$

- Lets calculate $a_{30}^{(1)}$

$$a_{30}^{(1)\prime} = a_{30}^{(1)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{30}^{(1)}}$$

$$a_{30}^{(1)\prime} = a_{30}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \frac{\partial (O_0^{(2)})}{\partial a_{30}^{(1)}}$$

$$a_{30}^{(1)\prime} = a_{30}^{(1)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)} x1 + a_{23}^{(0)} x2)}} \right)$$

- Lets calculate $a_{01}^{(0)}$

$$a_{01}^{(0)\prime} = a_{01}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{01}^{(0)}}$$

$$a_{01}^{(0)\prime} = a_{01}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial O_1^{(1)}} \cdot \frac{\partial (O_1^{(1)})}{\partial (a_{01}^{(0)})}$$

$$a_{01}^{(0)\prime} = a_{01}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)})$$

$$a_{01}^{(0)\prime} = a_{01}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)} x1 + a_{21}^{(0)} x2)}} \right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)} x1 + a_{21}^{(0)} x2)}} \right) \right)$$

- Lets calculate $a_{02}^{(0)}$

$$a_{02}^{(0)\prime} = a_{02}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{02}^{(0)}}$$

$$a_{02}^{(0)\prime} = a_{02}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial O_2^{(1)}} \cdot \frac{\partial (O_2^{(1)})}{\partial (a_{02}^{(0)})}$$

$$a_{02}^{(0)\prime} = a_{02}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)})$$

$$a_{02}^{(0)\prime} = a_{02}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)} x1 + a_{22}^{(0)} x2)}} \right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)} x1 + a_{22}^{(0)} x2)}} \right) \right)$$

- Lets calculate $a_{03}^{(0)}$

$$a_{03}^{(0)\prime} = a_{03}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{03}^{(0)}}$$

$$a_{03}^{(0)\prime} = a_{03}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial O_3^{(1)}} \cdot \frac{\partial (O_3^{(1)})}{\partial (a_{03}^{(0)})}$$

$$a_{03}^{(0)\prime} = a_{03}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)})$$

$$a_{03}^{(0)\prime} = a_{03}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)} x1 + a_{23}^{(0)} x2)}} \right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)} x1 + a_{23}^{(0)} x2)}} \right) \right)$$

- Lets calculate $a_{11}^{(0)}$

$$a_{11}^{(0)\prime} = a_{11}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{11}^{(0)}}$$

$$a_{11}^{(0)\prime} = a_{11}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial O_1^{(1)}} \cdot \frac{\partial (O_1^{(1)})}{\partial (a_{11}^{(0)})}$$

$$a_{11}^{(0)\prime} = a_{11}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot x1 \cdot O_1^{(1)} \cdot (1 - O_1^{(1)})$$

$$a_{11}^{(0)\prime} = a_{11}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot x1 \cdot \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)} x1 + a_{21}^{(0)} x2)}} \right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)} x1 + a_{21}^{(0)} x2)}} \right) \right)$$

- Lets calculate $a_{12}^{(0)}$

$$a_{12}^{(0)\prime} = a_{12}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{12}^{(0)}}$$

$$a_{12}^{(0)\prime} = a_{12}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial O_2^{(1)}} \cdot \frac{\partial (O_2^{(1)})}{\partial (a_{12}^{(0)})}$$

$$a_{12}^{(0)\prime} = a_{12}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot x1 \cdot O_2^{(1)} \cdot (1 - O_2^{(1)})$$

$$a_{12}^{(0)\prime} = a_{12}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot x1 \cdot \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)} x1 + a_{22}^{(0)} x2)}} \right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)} x1 + a_{22}^{(0)} x2)}} \right) \right)$$

- Lets calculate $a_{13}^{(0)}$

$$a_{13}^{(0)\prime} = a_{13}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial a_{13}^{(0)}}$$

$$a_{13}^{(0)\prime} = a_{13}^{(0)} - \alpha \frac{\partial (y - O_0^{(2)})^2}{\partial O_3^{(1)}} \cdot \frac{\partial (O_3^{(1)})}{\partial (a_{13}^{(0)})}$$

$$a_{13}^{(0)\prime} = a_{13}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot x1 \cdot O_3^{(1)} \cdot (1 - O_3^{(1)})$$

$$a_{13}^{(0)\prime} = a_{13}^{(0)} + 2 \cdot \alpha (y - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot x1 \cdot \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)} x1 + a_{23}^{(0)} x2)}} \right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)} x1 + a_{23}^{(0)} x2)}} \right) \right)$$

- Lets calculate $a_{21}^{(0)}$

$$a_{21}^{(0)\prime} = a_{21}^{(0)} - \alpha\frac{\partial(y - O_0^{(2)})^2}{\partial a_{21}^{(0)}}$$

$$a_{21}^{(0)\prime} = a_{21}^{(0)} - \alpha\frac{\partial(y - O_0^{(2)})^2}{\partial O_1^{(1)}} \cdot \frac{\partial(O_1^{(1)})}{\partial(a_{21}^{(0)})}$$

$$a_{21}^{(0)\prime} = a_{21}^{(0)} + 2 \cdot \alpha(y - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot x2 \cdot O_1^{(1)} \cdot (1 - O_1^{(1)})$$

$$a_{21}^{(0)\prime} = a_{21}^{(0)} + 2 \cdot \alpha(y - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot x2 \cdot \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)}x1 + a_{21}^{(0)}x2)}}\right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)}x1 + a_{21}^{(0)}x2)}}\right)\right)$$

- Lets calculate $a_{22}^{(0)}$

$$a_{22}^{(0)\prime} = a_{22}^{(0)} - \alpha\frac{\partial(y - O_0^{(2)})^2}{\partial a_{22}^{(0)}}$$

$$a_{22}^{(0)\prime} = a_{22}^{(0)} - \alpha\frac{\partial(y - O_0^{(2)})^2}{\partial O_2^{(1)}} \cdot \frac{\partial(O_2^{(1)})}{\partial(a_{22}^{(0)})}$$

$$a_{22}^{(0)\prime} = a_{22}^{(0)} + 2 \cdot \alpha(y - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot x2 \cdot O_2^{(1)} \cdot (1 - O_2^{(1)})$$

$$a_{22}^{(0)\prime} = a_{22}^{(0)} + 2 \cdot \alpha(y - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot x2 \cdot \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)}x1 + a_{22}^{(0)}x2)}}\right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)}x1 + a_{22}^{(0)}x2)}}\right)\right)$$

- Lets calculate $a_{23}^{(0)}$

$$a_{23}^{(0)\prime} = a_{23}^{(0)} - \alpha\frac{\partial(y - O_0^{(2)})^2}{\partial a_{23}^{(0)}}$$

$$a_{23}^{(0)\prime} = a_{23}^{(0)} - \alpha\frac{\partial(y - O_0^{(2)})^2}{\partial O_3^{(1)}} \cdot \frac{\partial(O_3^{(1)})}{\partial(a_{23}^{(0)})}$$

$$a_{23}^{(0)\prime} = a_{23}^{(0)} + 2 \cdot \alpha(y - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot x2 \cdot O_3^{(1)} \cdot (1 - O_3^{(1)})$$

$$a_{23}^{(0)\prime} = a_{23}^{(0)} + 2 \cdot \alpha(y - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot x2 \cdot \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)}x1 + a_{23}^{(0)}x2)}}\right) \cdot \left(1 - \left(\frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)}x1 + a_{23}^{(0)}x2)}}\right)\right)$$

Then lets solve the classification problem

In order to solve this problem in a clean manner, firstly lets calculate $O_1^{(1)}$, $O_2^{(1)}$ and $O_3^{(1)}$.

$$O_1^{(1)} = \sigma(a_{01}^{(0)} + a_{11}^{(0)}x1 + a_{21}^{(0)}x2) = \frac{1}{1 + e^{-(a_{01}^{(0)} + a_{11}^{(0)}x1 + a_{21}^{(0)}x2)}}$$
$$O_2^{(1)} = \sigma(a_{02}^{(0)} + a_{12}^{(0)}x1 + a_{22}^{(0)}x2) = \frac{1}{1 + e^{-(a_{02}^{(0)} + a_{12}^{(0)}x1 + a_{22}^{(0)}x2)}}$$
$$O_3^{(1)} = \sigma(a_{03}^{(0)} + a_{13}^{(0)}x1 + a_{23}^{(0)}x2) = \frac{1}{1 + e^{-(a_{03}^{(0)} + a_{13}^{(0)}x1 + a_{23}^{(0)}x2)}}$$

Then in order to calculate easily lets find $X_0^{(2)}$, $X_1^{(2)}$ and $X_2^{(2)}$.

$$X_0^{(2)} = O_0^{(1)} \cdot a_{00}^{(1)} + O_1^{(1)} \cdot a_{10}^{(1)} + O_2^{(1)} \cdot a_{20}^{(1)} + O_3^{(1)} \cdot a_{30}^{(1)}$$
$$X_1^{(2)} = O_0^{(1)} \cdot a_{01}^{(1)} + O_1^{(1)} \cdot a_{11}^{(1)} + O_2^{(1)} \cdot a_{21}^{(1)} + O_3^{(1)} \cdot a_{31}^{(1)}$$
$$X_2^{(2)} = O_0^{(1)} \cdot a_{02}^{(1)} + O_1^{(1)} \cdot a_{12}^{(1)} + O_2^{(1)} \cdot a_{22}^{(1)} + O_3^{(1)} \cdot a_{32}^{(1)}$$

Lets do some useful calculations.

$$\frac{dO_n^{(i)}}{dX_n^{(i)}} = O_n^{(i)} \cdot (1 - O_n^{(i)})$$
$$\frac{dO_k^{(i)}}{dX_n^{(i)}} = -O_n^{(i)} \cdot O_k^{(i)}$$

$$O_0^{(2)} = softmax(X_0^{(2)}, X^{(2)}) = \frac{e^{X_0^{(2)}}}{e^{X_0^{(2)}} + e^{X_1^{(2)}} + e^{X_2^{(2)}}}$$
$$O_1^{(2)} = softmax(X_1^{(2)}, X^{(2)}) = \frac{e^{X_0^{(2)}}}{e^{X_0^{(2)}} + e^{X_1^{(2)}} + e^{X_2^{(2)}}}$$
$$O_2^{(2)} = softmax(X_2^{(2)}, X^{(2)}) = \frac{e^{X_0^{(2)}}}{e^{X_0^{(2)}} + e^{X_1^{(2)}} + e^{X_2^{(2)}}}$$

Then now it is time to calculate all weight changes one by one.

– Lets calculate $a_{00}^{(1)}$

$$a_{00}^{(1)\prime} = a_{00}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{00}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{00}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{00}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{00}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{00}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot O_0^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{00}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{00}^{(1)}} = \frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)} \cdot O_0^{(2)} \cdot O_0^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{00}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{00}^{(1)}} = \frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)} \cdot O_0^{(2)} \cdot O_0^{(1)}$$

If we sum them up we get the final result.

$$a_{00}^{(1)} - \alpha \left(-l_0 \cdot (1 - O_0^2) + l_1 \cdot O_0^2 + l_2 \cdot O_0^2 \right)$$

– Lets calculate $a_{01}^{(1)}$

$$a_{01}^{(1)\prime} = a_{01}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{01}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{01}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{01}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{01}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{01}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_1^{(2)} \cdot O_0^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{01}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{01}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)} \cdot (1 - O_1^{(2)}) \cdot O_0^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{01}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{01}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_1^{(2)} \cdot O_0^{(1)}$$

If we sum them up we get the final result.

$$a_{01}^{(1)} - \alpha \left(l_0 \cdot O_1^{(2)} + l_1 \cdot (1 - O_1^{(2)}) - l_2 \cdot O_1^{(2)} \right)$$

– Lets calculate $a_{02}^{(1)}$

$$a_{02}^{(1)\prime} = a_{02}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{02}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{02}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{02}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{02}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{02}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_2^{(2)} \cdot O_0^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{02}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{02}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_2^{(2)} \cdot O_0^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{02}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{02}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)} \cdot (1 - O_2^{(2)}) \cdot O_0^{(1)}$$

If we sum them up we get the final result.

$$a_{02}^{(1)} - \alpha \left(l_0 \cdot O_2^{(2)} + l_1 \cdot O_2^{(2)} + l_2 \cdot (1 - O_2^{(2)}) \right)$$

– Lets calculate $a_{10}^{(1)}$

$$a_{10}^{(1)\prime} = a_{10}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{10}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{10}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{10}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{10}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{10}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)} (1 - O_0^{(2)}) \cdot O_1^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{10}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{10}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_0^{(2)} \cdot O_1^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{10}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{10}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_0^{(2)} \cdot O_1^{(1)}$$

If we sum them up we get the final result.

$$a_{10}^{(1)} - \alpha \left(-l_0 \cdot (1 - O_0^{(2)}) \cdot O_1^{(1)} + l_1 \cdot O_0^{(2)} \cdot O_1^{(1)} + l_2 \cdot O_0^{(2)} \cdot O_1^{(1)} \right)$$

– Lets calculate $a_{11}^{(1)}$

$$a_{11}^{(1)\prime} = a_{11}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{11}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{11}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{11}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{11}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{11}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_1^{(2)} \cdot O_1^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{11}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{11}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)} \cdot (1 - O_1^{(2)}) \cdot O_1^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{11}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{11}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_1^{(2)} \cdot O_1^{(1)}$$

If we sum them up we get the final result.

$$a_{11}^{(1)} - \alpha \left(l_0 \cdot O_1^{(2)} \cdot O_1^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot O_1^{(1)} + l_2 \cdot O_1^{(2)} \cdot O_1^{(1)} \right)$$

- Lets calculate $a_{12}^{(1)}$

$$a_{12}^{(1)\prime} = a_{12}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{12}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{12}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{12}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{12}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{12}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_2^{(2)} \cdot O_1^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{12}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{12}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_2^{(2)} \cdot O_1^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{12}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{12}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)} \cdot (1 - O_2^{(2)}) \cdot O_1^{(1)}$$

If we sum them up we get the final result.

$$a_{12}^{(1)} - \alpha \left(l_0 \cdot O_2^{(2)} \cdot O_1^{(1)} + l_1 \cdot O_2^{(2)} \cdot O_1^{(1)} - l_2 \cdot (1 - O_2^{(2)}) \cdot O_1^{(1)} \right)$$

- Lets calculate $a_{20}^{(1)}$

$$a_{20}^{(1)\prime} = a_{20}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{20}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{20}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{20}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{20}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{20}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)} (1 - O_0^{(2)}) \cdot O_2^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{20}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{20}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_0^{(2)} \cdot O_2^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{20}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{20}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_0^{(2)} \cdot O_2^{(1)}$$

If we sum them up we get the final result.

$$a_{20}^{(1)} - \alpha \left(-l_0 \cdot (1 - O_0^{(2)}) \cdot O_2^{(1)} + l_1 \cdot O_0^{(2)} \cdot O_2^{(1)} + l_2 \cdot O_0^{(2)} \cdot O_2^{(1)} \right)$$

- Lets calculate $a_{21}^{(1)}$

$$a_{21}^{(1)\prime} = a_{21}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{21}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{21}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{21}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{21}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{21}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_1^{(2)} \cdot O_2^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{21}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{21}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)} \cdot (1 - O_1^{(2)}) \cdot O_2^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{21}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{21}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_1^{(2)} \cdot O_2^{(1)}$$

If we sum them up we get the final result.

$$a_{21}^{(1)} - \alpha \left(l_0 \cdot O_1^{(2)} \cdot O_2^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot O_2^{(1)} - l_2 \cdot O_1^{(2)} \cdot O_2^{(1)} \right)$$

– Lets calculate $a_{22}^{(1)}$

$$a_{22}^{(1)\prime} = a_{22}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{22}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{22}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{22}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{22}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{22}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_2^{(2)} \cdot O_2^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{22}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{22}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_2^{(2)} \cdot O_2^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{22}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{22}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)} \cdot (1 - O_2^{(2)}) \cdot O_2^{(1)}$$

If we sum them up we get the final result.

$$a_{22}^{(1)} - \alpha \left(l_0 \cdot O_2^{(2)} \cdot O_2^{(1)} + l_1 \cdot O_2^{(2)} \cdot O_2^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot O_2^{(1)} \right)$$

– Lets calculate $a_{30}^{(1)}$

$$a_{30}^{(1)\prime} = a_{30}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{30}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{30}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{30}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{30}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{30}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)} (1 - O_0^{(2)}) \cdot O_3^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{30}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{30}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_0^{(2)} \cdot O_3^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{30}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial a_{30}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_0^{(2)} \cdot O_3^{(1)}$$

If we sum them up we get the final result.

$$a_{30}^{(1)} - \alpha \left(-l_0 \cdot (1 - O_0^{(2)}) \cdot O_3^{(1)} + l_1 \cdot O_0^{(2)} \cdot O_3^{(1)} + l_2 \cdot O_0^{(2)} \cdot O_3^{(1)} \right)$$

– Lets calculate $a_{31}^{(1)}$

$$a_{31}^{(1)\prime} = a_{31}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{31}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{31}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{31}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{31}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{31}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_1^{(2)} \cdot O_3^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{31}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{31}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)} \cdot (1 - O_1^{(2)}) \cdot O_3^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{31}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial a_{31}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot -O_2^{(2)} \cdot O_1^{(2)} \cdot O_3^{(1)}$$

If we sum them up we get the final result.

$$a_{31}^{(1)} - \alpha \left(l_0 \cdot O_1^{(2)} \cdot O_3^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot O_3^{(1)} - l_2 \cdot O_1^{(2)} \cdot O_3^{(1)} \right)$$

– Lets calculate $a_{32}^{(1)}$

$$a_{32}^{(1)\prime} = a_{32}^{(1)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{32}^{(1)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{32}^{(1)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{32}^{(1)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{32}^{(1)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{32}^{(1)}} = -\frac{l_0}{O_0^{(2)}} \cdot -O_0^{(2)} O_2^{(2)} \cdot O_3^{(1)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{32}^{(1)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{32}^{(1)}} = -\frac{l_1}{O_1^{(2)}} \cdot -O_1^{(2)} \cdot O_2^{(2)} \cdot O_3^{(1)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{32}^{(1)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial a_{32}^{(1)}} = -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)} \cdot (1 - O_2^{(2)}) \cdot O_3^{(1)}$$

If we sum them up we get the final result.

$$a_{32}^{(1)} - \alpha \left(l_0 \cdot O_2^{(2)} \cdot O_3^{(1)} + l_1 \cdot O_2^{(2)} \cdot O_3^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot O_3^{(1)} \right)$$

– Lets calculate $a_{01}^{(0)}$

$$a_{01}^{(0)\prime} = a_{01}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{01}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{01}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{01}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$\begin{aligned} -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{01}^{(0)}} &= -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{01}^{(0)}} \\ &= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_0^{(0)} \end{aligned}$$

$$\begin{aligned} -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{01}^{(0)}} &= -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{01}^{(0)}} \\ &= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1 - O_1^{(2)}) \cdot a_{11}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_0^{(0)} \end{aligned}$$

$$\begin{aligned} -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{01}^{(0)}} &= -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{01}^{(0)}} \\ &= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{12}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_0^{(0)} \end{aligned}$$

If we sum them up we get the final result.

$$a_{01}^{(0)} - \alpha \cdot (O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_0^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{10}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{11}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{12}^{(1)} \right)$$

– Lets calculate $a_{02}^{(0)}$

$$a_{02}^{(0)\prime} = a_{02}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{02}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{02}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{02}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$\begin{aligned} -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{02}^{(0)}} &= -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{02}^{(0)}} \\ &= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_0^{(0)} \end{aligned}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{02}^{(0)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{02}^{(0)}}$$
$$= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1-O_1^{(2)}) \cdot a_{21}^{(1)} \cdot O_2^{(1)} \cdot (1-O_2^{(1)}) \cdot O_0^{(0)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{02}^{(0)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{02}^{(0)}}$$
$$= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1-O_2^{(2)}) \cdot a_{22}^{(1)} \cdot O_2^{(1)} \cdot (1-O_2^{(1)}) \cdot O_0^{(0)}$$

If we sum them up we get the final result.

$$a_{02}^{(0)} - \alpha \cdot (O_2^{(1)} \cdot (1-O_2^{(1)}) \cdot O_0^{(0)}) \cdot \left(l_0 \cdot (1-O_0^{(2)}) \cdot a_{20}^{(1)} + l_1 \cdot (1-O_1^{(2)}) \cdot a_{21}^{(1)} + l_2 \cdot (1-O_2^{(2)}) \cdot a_{22}^{(1)}\right)$$

- Lets calculate $a_{03}^{(0)}$

$$a_{03}^{(0)\prime} = a_{03}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{03}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{03}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{03}^{(0)}}\right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{03}^{(0)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{03}^{(0)}}$$
$$= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1-O_0^{(2)}) \cdot a_{30}^{(1)} \cdot O_3^{(1)} \cdot (1-O_3^{(1)}) \cdot O_0^{(0)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{03}^{(0)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{03}^{(0)}}$$
$$= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1-O_1^{(2)}) \cdot a_{31}^{(1)} \cdot O_3^{(1)} \cdot (1-O_3^{(1)}) \cdot O_0^{(0)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{03}^{(0)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{03}^{(0)}}$$
$$= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1-O_2^{(2)}) \cdot a_{32}^{(1)} \cdot O_3^{(1)} \cdot (1-O_3^{(1)}) \cdot O_0^{(0)}$$

If we sum them up we get the final result.

$$a_{03}^{(0)} - \alpha \cdot (O_3^{(1)} \cdot (1-O_3^{(1)}) \cdot O_0^{(0)}) \cdot \left(l_0 \cdot (1-O_0^{(2)}) \cdot a_{30}^{(1)} + l_1 \cdot (1-O_1^{(2)}) \cdot a_{31}^{(1)} + l_2 \cdot (1-O_2^{(2)}) \cdot a_{32}^{(1)}\right)$$

- Lets calculate $a_{11}^{(0)}$

$$a_{11}^{(0)\prime} = a_{11}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{11}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{11}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{11}^{(0)}}\right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{01}^{(0)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{11}^{(0)}}$$
$$= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1-O_0^{(2)}) \cdot a_{10}^{(1)} \cdot O_1^{(1)} \cdot (1-O_1^{(1)}) \cdot O_1^{(0)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{01}^{(0)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{11}^{(0)}}$$
$$= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1-O_1^{(2)}) \cdot a_{11}^{(1)} \cdot O_1^{(1)} \cdot (1-O_1^{(1)}) \cdot O_1^{(0)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{01}^{(0)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{11}^{(0)}}$$
$$= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{12}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_1^{(0)}$$

If we sum them up we get the final result.

$$a_{11}^{(0)} - \alpha \cdot (O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_1^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{10}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{11}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{12}^{(1)}\right)$$

– Lets calculate $a_{12}^{(0)}$

$$a_{12}^{(0)\prime} = a_{12}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{12}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{12}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{12}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{12}^{(0)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{12}^{(0)}}$$
$$= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_1^{(0)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{12}^{(0)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{12}^{(0)}}$$
$$= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1 - O_1^{(2)}) \cdot a_{21}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_1^{(0)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{12}^{(0)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{12}^{(0)}}$$
$$= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{22}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_1^{(0)}$$

If we sum them up we get the final result.

$$a_{12}^{(0)} - \alpha \cdot (O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_1^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{20}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{21}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{22}^{(1)}\right)$$

– Lets calculate $a_{13}^{(0)}$

$$a_{13}^{(0)\prime} = a_{13}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{13}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{13}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{13}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{13}^{(0)}} = -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{13}^{(0)}}$$
$$= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_1^{(0)}$$

$$-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{13}^{(0)}} = -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{13}^{(0)}}$$
$$= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1 - O_1^{(2)}) \cdot a_{31}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_1^{(0)}$$

$$-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{13}^{(0)}} = -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{13}^{(0)}}$$
$$= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{32}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_1^{(0)}$$

If we sum them up we get the final result.

$$a_{13}^{(0)} - \alpha \cdot (O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_1^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{30}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{31}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{32}^{(1)}\right)$$

– Lets calculate $a_{21}^{(0)}$

$$a_{21}^{(0)\prime} = a_{21}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{21}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{21}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{21}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$\begin{aligned} -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{21}^{(0)}} &= -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{21}^{(0)}} \\ &= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{10}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_2^{(0)} \end{aligned}$$

$$\begin{aligned} -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{21}^{(0)}} &= -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{21}^{(0)}} \\ &= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1 - O_1^{(2)}) \cdot a_{11}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_2^{(0)} \end{aligned}$$

$$\begin{aligned} -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{21}^{(0)}} &= -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_1^{(1)}} \cdot \frac{\partial O_1^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{21}^{(0)}} \\ &= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{12}^{(1)} \cdot O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_2^{(0)} \end{aligned}$$

If we sum them up we get the final result.

$$a_{21}^{(0)} - \alpha \cdot (O_1^{(1)} \cdot (1 - O_1^{(1)}) \cdot O_2^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{10}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{11}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{12}^{(1)} \right)$$

– Lets calculate $a_{22}^{(0)}$

$$a_{22}^{(0)\prime} = a_{22}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{22}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{22}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{22}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$\begin{aligned} -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{22}^{(0)}} &= -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{22}^{(0)}} \\ &= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{20}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_2^{(0)} \end{aligned}$$

$$\begin{aligned} -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{22}^{(0)}} &= -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{22}^{(0)}} \\ &= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1 - O_1^{(2)}) \cdot a_{21}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_2^{(0)} \end{aligned}$$

$$\begin{aligned} -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{22}^{(0)}} &= -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_2^{(1)}} \cdot \frac{\partial O_2^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{22}^{(0)}} \\ &= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{22}^{(1)} \cdot O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_2^{(0)} \end{aligned}$$

If we sum them up we get the final result.

$$a_{22}^{(0)} - \alpha \cdot (O_2^{(1)} \cdot (1 - O_2^{(1)}) \cdot O_2^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{20}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{21}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{22}^{(1)} \right)$$

– Lets calculate $a_{23}^{(0)}$

$$a_{23}^{(0)\prime} = a_{13}^{(0)} - \alpha \cdot \left(-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{23}^{(0)}} - \frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{23}^{(0)}} - \frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{23}^{(0)}} \right)$$

Lets calculate each derivative one by one and then sum them up.

$$
\begin{aligned}
-\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial a_{23}^{(0)}} &= -\frac{\partial l_0 \cdot \log(O_0^{(2)})}{\partial O_0^{(2)}} \cdot \frac{\partial O_0^{(2)}}{\partial X_0^{(2)}} \cdot \frac{\partial X_0^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{23}^{(0)}} \\
&= -\frac{l_0}{O_0^{(2)}} \cdot O_0^{(2)}(1 - O_0^{(2)}) \cdot a_{30}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_2^{(0)}
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial a_{23}^{(0)}} &= -\frac{\partial l_1 \cdot \log(O_1^{(2)})}{\partial O_1^{(2)}} \cdot \frac{\partial O_1^{(2)}}{\partial X_1^{(2)}} \cdot \frac{\partial X_1^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{23}^{(0)}} \\
&= -\frac{l_1}{O_1^{(2)}} \cdot O_1^{(2)}(1 - O_1^{(2)}) \cdot a_{31}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_2^{(0)}
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial a_{23}^{(0)}} &= -\frac{\partial l_2 \cdot \log(O_2^{(2)})}{\partial O_2^{(2)}} \cdot \frac{\partial O_2^{(2)}}{\partial X_2^{(2)}} \cdot \frac{\partial X_2^{(2)}}{\partial O_3^{(1)}} \cdot \frac{\partial O_3^{(1)}}{\partial z} \cdot \frac{\partial z}{\partial a_{23}^{(0)}} \\
&= -\frac{l_2}{O_2^{(2)}} \cdot O_2^{(2)}(1 - O_2^{(2)}) \cdot a_{32}^{(1)} \cdot O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_2^{(0)}
\end{aligned}
$$

If we sum them up we get the final result.

$$
a_{23}^{(0)} - \alpha \cdot (O_3^{(1)} \cdot (1 - O_3^{(1)}) \cdot O_2^{(0)}) \cdot \left(l_0 \cdot (1 - O_0^{(2)}) \cdot a_{30}^{(1)} + l_1 \cdot (1 - O_1^{(2)}) \cdot a_{31}^{(1)} + l_2 \cdot (1 - O_2^{(2)}) \cdot a_{32}^{(1)}\right)
$$