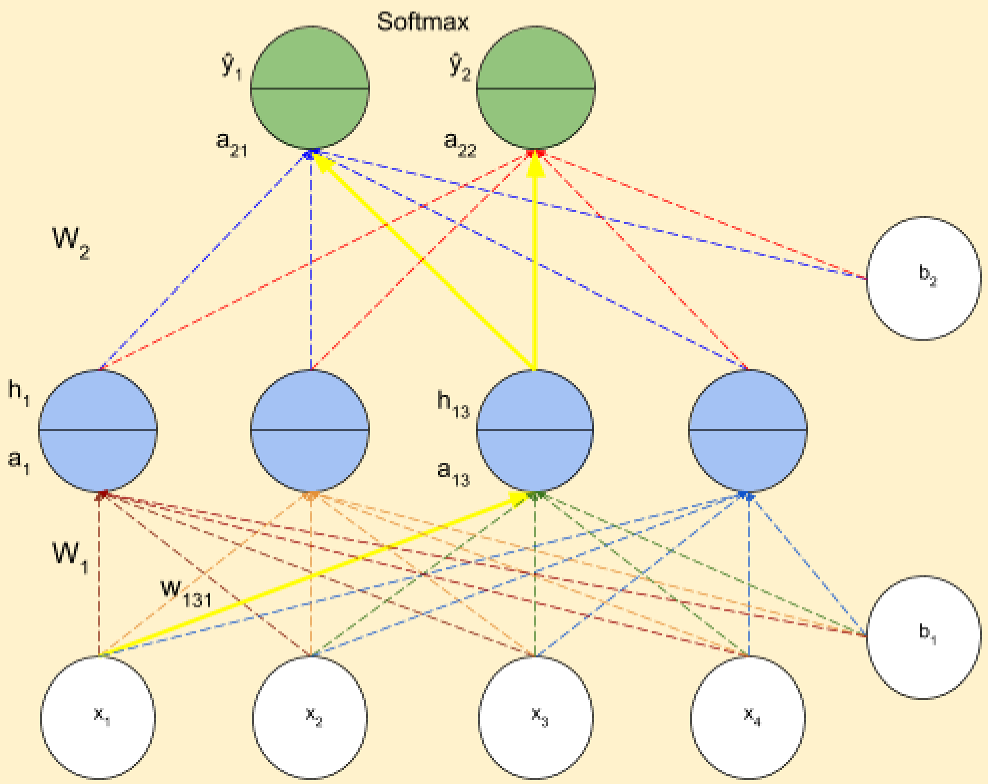
PadhAI: Backpropagation - the light math version

One Fourth Labs

Multiple Paths

Can we see one more example?

1. Let's look at a different weight from the previous example, which would require multiple paths to perform the calculations



2. Here are the parameters of the network
   a. $b = [\,0 \;\; 0\,]$
   b.

$$W_1 = \begin{bmatrix} 0.1 & 0.3 & 0.8 & -0.4 \\ -0.3 & -0.2 & 0.5 & 0.5 \\ -0.3 & 0 & 0.5 & 0.4 \\ 0.2 & 0.5 & -0.9 & 0.7 \end{bmatrix}$$

c.

$$W_2 = \begin{bmatrix} 0.5 & 0.8 & 0.2 & 0.4 \\ 0.5 & 0.2 & 0.3 & -0.5 \end{bmatrix}$$

d. x = [2 5 3 3] true distribution y = [1 0]

3. Now, we want to find the partial derivative w.r.t w_{212} as highlighted in the figure $\frac{\partial L}{\partial w_{212}}$
4. $\frac{\partial L}{\partial w_{131}} = (\frac{\partial L}{\partial a_{13}}).(\frac{\partial a_{13}}{\partial w_{131}}) = (\frac{\partial L}{\partial h_{13}}).(\frac{\partial h_{13}}{\partial a_{13}}).(\frac{\partial a_{13}}{\partial w_{131}}) = (\frac{\partial L}{\partial a_{21}}.\frac{\partial a_{21}}{\partial h_{13}} + \frac{\partial L}{\partial a_{22}}.\frac{\partial a_{22}}{\partial h_{13}}).(\frac{\partial h_{13}}{\partial a_{13}}).(\frac{\partial a_{13}}{\partial w_{131}})$ a
5. The final split is $\frac{\partial L}{\partial w_{131}} = (\frac{\partial L}{\partial \hat{y}_1}.\frac{\partial \hat{y}_1}{\partial a_{21}}.\frac{\partial a_{21}}{\partial h_{13}} + \frac{\partial L}{\partial \hat{y}_2}.\frac{\partial \hat{y}_2}{\partial a_{22}}.\frac{\partial a_{22}}{\partial h_{13}}).(\frac{\partial h_{13}}{\partial a_{13}}).(\frac{\partial a_{13}}{\partial w_{131}})$
6. Let us sequentially solve both splits

$\frac{\partial L}{\partial \hat{y}_1} = -2(y_1 - \hat{y}_1) = -0.46$	$\frac{\partial L}{\partial \hat{y}_2} = -2(y_2 - \hat{y}_2) = 0.46$
$\frac{\partial \hat{y}_1}{\partial a_{21}} = \hat{y}_1(1 - \hat{y}_1) = 0.1771$	$\frac{\partial \hat{y}_2}{\partial a_{22}} = \hat{y}_2(1 - \hat{y}_2) = 0.1771$
$\frac{\partial a_{21}}{\partial h_{13}} = w_{213} = 0.2$	$\frac{\partial a_{22}}{\partial h_{13}} = w_{223} = 0.3$
$\frac{\partial h_{13}}{\partial a_{13}} = h_{13} * (1 - h_{13}) = 0.0979$	$\frac{\partial h_{13}}{\partial a_{13}} = h_{13} * (1 - h_{13}) = 0.0979$
$\frac{\partial a_{13}}{\partial w_{131}} = x_1 = 2$	$\frac{\partial a_{13}}{\partial w_{131}} = x_1 = 2$
Path1: $(-0.46 * 0.1771 * 0.2 * 0.0979 * 2) = -0.003190$	Path1: $(0.46 * 0.1771 * 0.3 * 0.0979 * 2) = 0.004785$
Sum of the paths is $\frac{\partial L}{\partial w_{131}} = 0.001595$	

7. Now we can calculate the updated value of w_{212}
8. $w_{131} = w_{131} - \eta(\frac{\partial L}{\partial w_{131}})$
   a. $w_{131} = -0.3 - (1) * (0.001595)$
   b. $w_{131} = -0.301595$
9. We can repeat this process for each weight