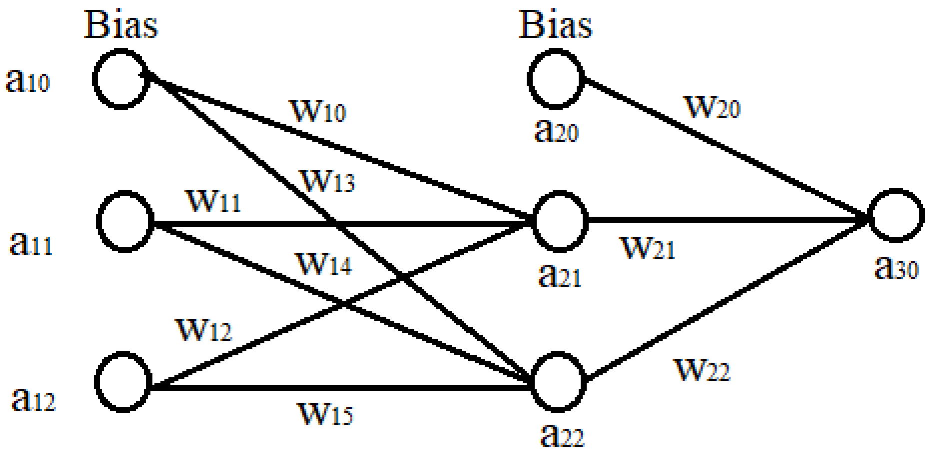
Backpropagation Algorithm for Neural Network Construction



Let's assume the input layer of the neural network is:

$$a_1 = \begin{bmatrix} a_{10} \\ a_{11} \\ a_{12} \end{bmatrix}; \text{ Where } a_{10} \text{ is Bias.}$$

And the weight associated to layer 1 & 2 is:

$$w_1 = \begin{bmatrix} w_{10} & w_{13} \\ w_{11} & w_{14} \\ w_{12} & w_{15} \end{bmatrix}$$

So, the second layer of the neural network will be:

$$a_2 = \begin{bmatrix} 1 \\ \left(1 + e^{-(a_{10}w_{10} + a_{11}w_{11} + a_{12}w_{12})}\right)^{-1} \\ \left(1 + e^{-(a_{10}w_{13} + a_{11}w_{14} + a_{12}w_{15})}\right)^{-1} \end{bmatrix} = \begin{bmatrix} 1 \\ (1 + e^{-z_{21}})^{-1} \\ (1 + e^{-z_{22}})^{-1} \end{bmatrix} = \begin{bmatrix} a_{20} \\ a_{21} \\ a_{22} \end{bmatrix}; \text{ Where } a_{20} \text{ is Bias.}$$

Again, assume the weight associated to layer 2 & 3 is:

$$w_1 = \begin{bmatrix} w_{20} \\ w_{21} \\ w_{22} \end{bmatrix}$$

And the final layer is:

$$a_{30} = \left(1 + e^{-(a_{20}w_{20} + a_{21}w_{21} + a_{22}w_{22})}\right)^{-1} = (1 + e^{-z_{30}})^{-1}$$

Total error in the feedforward path, $\boldsymbol{E = \frac{1}{2}(a_{30} - y)^2}$ .Where 'y' is the desired output.

Backward Path Calculation:

$$\boldsymbol{\frac{\partial E}{\partial w_{20}} = \frac{\partial}{\partial w_{20}}\left(\frac{1}{2}(a_{30} - y)^2\right) = (a_{30} - y)\frac{\partial}{\partial w_{20}}a_{30} = a_{20}(a_{30} - y)\,(a_{30})^2 e^{-z_{30}} = k_{20}}$$

Similarly, $\boldsymbol{\frac{\partial E}{\partial w_{21}} = a_{21}(a_{30} - y)\,(a_{30})^2 e^{-z_{30}} = k_{21}}$

$$\boldsymbol{\frac{\partial E}{\partial w_{22}} = a_{22}(a_{30} - y)\,(a_{30})^2 e^{-z_{30}} = k_{22}}$$

$$\text{So, } \frac{\partial E}{\partial w_2} = (a_{30} - y)\,(a_{30})^2 e^{-z_{30}} \begin{bmatrix} a_{20} \\ a_{21} \\ a_{22} \end{bmatrix} = (a_{30} - y)\,(a_{30})^2 e^{-z_{30}} a_2 = \begin{bmatrix} k_{20} \\ k_{21} \\ k_{22} \end{bmatrix}$$

Applying chain rule:

$$\frac{\partial E}{\partial w_{10}} = \frac{\partial E}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial w_{10}} \qquad \frac{\partial E}{\partial w_{13}} = \frac{\partial E}{\partial a_{22}} \cdot \frac{\partial a_{22}}{\partial z_{22}} \cdot \frac{\partial z_{22}}{\partial w_{13}}$$

$$\frac{\partial E}{\partial w_{11}} = \frac{\partial E}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial w_{11}} \qquad \frac{\partial E}{\partial w_{14}} = \frac{\partial E}{\partial a_{22}} \cdot \frac{\partial a_{22}}{\partial z_{22}} \cdot \frac{\partial z_{22}}{\partial w_{14}}$$

$$\frac{\partial E}{\partial w_{12}} = \frac{\partial E}{\partial a_{21}} \cdot \frac{\partial a_{21}}{\partial z_{21}} \cdot \frac{\partial z_{21}}{\partial w_{12}} \qquad \frac{\partial E}{\partial w_{15}} = \frac{\partial E}{\partial a_{22}} \cdot \frac{\partial a_{22}}{\partial z_{22}} \cdot \frac{\partial z_{22}}{\partial w_{15}}$$

$$\therefore \frac{\partial E}{\partial a_{21}} = \frac{\partial}{\partial a_{21}} \frac{1}{2} (a_{30} - y)^2$$

$$\Rightarrow \frac{\partial E}{\partial a_{21}} = (a_{30} - y) \frac{\partial}{\partial a_{21}} (a_{30})$$

$$\Rightarrow \frac{\partial E}{\partial a_{21}} = (a_{30} - y) \frac{\partial}{\partial a_{21}} (1 + e^{-z_{30}})^{-1}$$

$$\Rightarrow \frac{\partial E}{\partial a_{21}} = -(a_{30} - y)(1 + e^{-z_{30}})^{-2} \frac{\partial}{\partial a_{21}} e^{-z_{30}}$$

$$\Rightarrow \frac{\partial E}{\partial a_{21}} = (a_{30} - y)(a_{30})^2 e^{-z_{30}} \frac{\partial}{\partial a_{21}} z_{30}$$

$$\Rightarrow \frac{\partial E}{\partial a_{21}} = (a_{30} - y)(a_{30})^2 e^{-z_{30}} \frac{\partial}{\partial a_{21}} (a_{20} w_{20} + a_{21} w_{21} + a_{22} w_{22})$$

$$\textit{i.e. } \frac{\partial E}{\partial a_{21}} = (a_{30} - y)(a_{30})^2 e^{-z_{30}} w_{21}$$

$$\text{Similarly, } \frac{\partial E}{\partial a_{22}} = (a_{30} - y)(a_{30})^2 e^{-z_{30}} w_{22}$$

$$\text{And, } \frac{\partial a_{21}}{\partial z_{21}} = \frac{\partial}{\partial z_{21}} (1 + e^{-z_{21}})^{-1}$$

$$\Rightarrow \frac{\partial a_{21}}{\partial z_{21}} = (1 + e^{-z_{21}})^{-2} e^{-z_{21}} \frac{\partial}{\partial z_{21}} (z_{21})$$

$$\textit{i.e. } \frac{\partial a_{21}}{\partial z_{21}} = (a_{21})^2 e^{-z_{21}}$$

$$\text{Similarly, } \frac{\partial a_{22}}{\partial z_{22}} = (a_{22})^2 e^{-z_{22}}$$

$$\text{And, } \frac{\partial z_{21}}{\partial w_{10}} = \frac{\partial}{\partial w_{10}} (a_{10} w_{10} + a_{11} w_{11} + a_{12} w_{12}) = a_{10}$$

$$\frac{\partial z_{21}}{\partial w_{11}} = \frac{\partial}{\partial w_{11}} (a_{10} w_{10} + a_{11} w_{11} + a_{12} w_{12}) = a_{11}$$

$$\frac{\partial z_{21}}{\partial w_{12}} = \frac{\partial}{\partial w_{12}} (a_{10} w_{10} + a_{11} w_{11} + a_{12} w_{12}) = a_{12}$$

$$\frac{\partial z_{22}}{\partial w_{13}} = \frac{\partial}{\partial w_{13}} (a_{10} w_{13} + a_{11} w_{14} + a_{12} w_{15}) = a_{10}$$

$$\frac{\partial z_{22}}{\partial w_{14}} = \frac{\partial}{\partial w_{14}} (a_{10} w_{13} + a_{11} w_{14} + a_{12} w_{15}) = a_{11}$$

$$\frac{\partial z_{22}}{\partial w_{15}} = \frac{\partial}{\partial w_{15}} (a_{10} w_{13} + a_{11} w_{14} + a_{12} w_{15}) = a_{12}$$

$$\therefore \begin{bmatrix} \frac{\partial E}{\partial w_{10}} \\ \frac{\partial E}{\partial w_{11}} \\ \frac{\partial E}{\partial w_{12}} \end{bmatrix} = (a_{30} - y)(a_{30} \cdot a_{21})^2 \cdot e^{-(z_{30}+z_{21})} \cdot w_{21} \cdot a_1$$

$$\&, \begin{bmatrix} \frac{\partial E}{\partial w_{13}} \\ \frac{\partial E}{\partial w_{14}} \\ \frac{\partial E}{\partial w_{15}} \end{bmatrix} = (a_{30} - y)(a_{30} \cdot a_{22})^2 \cdot e^{-(z_{30}+z_{22})} \cdot w_{22} \cdot a_1$$

Now, the updated weight will be:

$$\begin{bmatrix} w_{10} & w_{13} \\ w_{11} & w_{14} \\ w_{12} & w_{15} \end{bmatrix} = \begin{bmatrix} w_{10} & w_{13} \\ w_{11} & w_{14} \\ w_{12} & w_{15} \end{bmatrix} - \alpha \begin{bmatrix} \frac{\partial E}{\partial w_{10}} & \frac{\partial E}{\partial w_{13}} \\ \frac{\partial E}{\partial w_{11}} & \frac{\partial E}{\partial w_{14}} \\ \frac{\partial E}{\partial w_{12}} & \frac{\partial E}{\partial w_{15}} \end{bmatrix}$$

$$\text{And, } \begin{bmatrix} w_{20} \\ w_{21} \\ w_{22} \end{bmatrix} = \begin{bmatrix} w_{20} \\ w_{21} \\ w_{22} \end{bmatrix} - \alpha \begin{bmatrix} \frac{\partial E}{\partial w_{20}} \\ \frac{\partial E}{\partial w_{21}} \\ \frac{\partial E}{\partial w_{22}} \end{bmatrix}$$