Given a matrix of weights: W = $\begin{bmatrix} W_{0,0} & \cdots & W_{0,k} \\ \vdots & \ddots & \vdots \\ W_{m,0} & \cdots & W_{m,k} \end{bmatrix}$

An input vector: x = $\begin{bmatrix} x_0 \\ \vdots \\ x_k \end{bmatrix}$ A bias vector: b = $\begin{bmatrix} b_0 \\ \vdots \\ b_m \end{bmatrix}$ And a target vector of desired outputs: T = $\begin{bmatrix} T_0 \\ \vdots \\ T_m \end{bmatrix}$

$$\boldsymbol{Weighted\ Sum\ of\ Output\ Neuron\ i = \ S_i = b_i + \sum_{j=0}^{k} W_{i,j} * x_j}$$

$$\boldsymbol{Weighted\ Sum\ of\ all\ Output\ Neurons\ (vector) = S = W * x + b}$$

$$\boldsymbol{Sigmoid\ Output\ of\ Output\ Neuron\ i = \ O_i = \frac{1}{1 + e^{-S_i}} = (1 + e^{-S_i})^{-1}}$$

$$\boldsymbol{Sigmoid\ Output\ of\ all\ Output\ Neurons\ (vector) = O = } \begin{bmatrix} O_0 \\ \vdots \\ O_m \end{bmatrix}$$

$$\boldsymbol{Mean\ Squared\ Error\ of\ all\ Output\ Neurons = MSE = \sum_{n=0}^{m} (O_n - T_n)^2}$$

$$\boldsymbol{\frac{\partial MSE}{\partial O_i} = 2(O_n - T_n)}$$

$$\boldsymbol{\frac{\partial O_i}{\partial S_i} = O_i(1 - O_i) = \frac{1}{1 + e^{-S_i}} * (1 - \frac{1}{1 + e^{-S_i}})}$$

$$\boldsymbol{\frac{\partial S_i}{\partial W_{i,j}} = X_j}$$

$$\boldsymbol{\frac{\partial MSE}{\partial W_{i,j}} = X_j * 2(O_i - T_i) * O_i(1 - O_i)}$$