Error function:-

$$E = \frac{1}{2}\sum_{i=1}^{M}\left(y^{(i)} - \hat{y}^{(i)}\right)^2$$

Applying sigmoid function to both the layers

$$\hat{y} = \text{Sigmoid}(v^T z)$$

$$z_j = \text{Sigmoid}(w_j^T x)$$

Taking Partial Derivative of Error function using chain Rule

$$\frac{\partial E}{\partial v} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v} \quad \& \quad \frac{\partial E}{\partial w_j} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_j}$$

So,

$$\frac{\partial E}{\partial v} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial v}$$

$$= \frac{1}{\cancel{2}} \times \cancel{2} \sum_{i=1}^{M}\left(y^{(i)} - \hat{y}^{(i)}\right)\hat{y}^{(i)} \cdot \left(1-\hat{y}^{(i)}\right) z_j^{(i)}$$

$$\therefore \frac{\partial E}{\partial v} = \sum_{i=1}^{M}\left(y^{(i)} - \hat{y}^{(i)}\right)\hat{y}^{(i)} \cdot \left(1-\hat{y}^{(i)}\right) z_j^{(i)}$$

Update Equation for output layer,

$$v_j \leftarrow v_j - \eta \sum_{i=1}^{M}\left(y^{(i)} - \hat{y}^{(i)}\right)\hat{y}^{(i)}\left(1-\hat{y}^{(i)}\right) z_j^{(i)}$$

And,

$$\frac{\partial E}{\partial w_j} = \frac{\partial E}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_j}$$

$$= \frac{1}{\cancel{2}} \times \cancel{2} \sum_{i=1}^{M}\left(y^{(i)} - \hat{y}^{(i)}\right) \cdot \hat{y}^{(i)}\left(1-\hat{y}^{(i)}\right) v_j z_j^{(i)}\left(1-z_j^{(i)}\right) x^{(i)}$$

$$w_j \leftarrow w_j - \eta \sum_{i=1}^{M}\left(y^{(i)} - \hat{y}^{(i)}\right)\hat{y}^{(i)}\left(1-\hat{y}^{(i)}\right) v_j z_j^{(i)}\left(1-z_j^{(i)}\right) x^{(i)}$$

Comparison :-

Loglikelihood

$$v_j \leftarrow v_j - \eta \sum_{i=1}^{m} (\hat{y}^{(i)} - y^{(i)}) z^{(i)}$$

$$w_j \leftarrow w_j - \eta \sum_{i=1}^{m} (\hat{y}^{(i)} - y^{(i)}) v_j z_j^{(i)} (1 - z_j^{(i)}) x^{(i)}$$

The term for $\hat{y}$ uses sigmoid function for activation but it cancels out in loglikelihood whereas it remains in error method using mean square