$$\mathbf{a}^{(2)} = \sigma(\mathbf{z}^{(2)}) = \sigma\left(W_{1,1}^{(2)}\sigma\left(W_{1,1}^{(1)}a_1^{(0)} + b_1^{(1)}\right) + W_{1,2}^{(2)}\sigma\left(W_{2,1}^{(1)}a_1^{(0)} + b_2^{(1)}\right) + b_1^{(2)}\right) =$$

$$= \sigma\left(W_{1,1}^{(2)}\sigma\left(W_{1,1}^{(1)}x + b_1^{(1)}\right) + W_{1,2}^{(2)}\sigma\left(W_{2,1}^{(1)}x + b_2^{(1)}\right) + b_1^{(2)}\right) =$$

$$= \sigma\left(W_{1,1}^{(2)}\sigma\left(z_1^{(1)}\right) + W_{1,2}^{(2)}\sigma\left(z_2^{(1)}\right) + b_1^{(2)}\right) = y$$

The cost function is defined as

$$C(\mathbf{x}, \mathbf{y}) = \frac{1}{2}\sum_{i=1}^{n}[y_i - f(x_i)]^2 = \tag{1}$$

$$\tag{2}$$

$$= \frac{1}{2}\sum_{i=1}^{n}\left[\sigma\left(W_{1,1}^{(2)}\sigma\left(W_{1,1}^{(1)}x_i + b_1^{(1)}\right) + W_{1,2}^{(2)}\sigma\left(W_{2,1}^{(1)}x_i + b_2^{(1)}\right) + b_1^{(2)}\right) - f(x_i)\right]^2 = \tag{3}$$

$$\tag{4}$$

$$= C(W_{1,1}^{(1)}, W_{2,1}^{(1)}, b_1^{(1)}, b_2^{(1)}, W_{1,1}^{(2)}, W_{1,2}^{(2)}, b_1^{(2)}, \mathbf{x}) \tag{5}$$

$$W_{1,1}^{(1)} \to W_{1,1}^{(1)} - h\frac{\partial C}{\partial W_{1,1}^{(1)}} \tag{6}$$

$$W_{2,1}^{(1)} \to W_{2,1}^{(1)} - h\frac{\partial C}{\partial W_{2,1}^{(1)}} \tag{7}$$

$$W_{1,1}^{(2)} \to W_{1,1}^{(2)} - h\frac{\partial C}{\partial W_{1,1}^{(2)}} \tag{8}$$

$$W_{1,2}^{(2)} \to W_{1,2}^{(2)} - h\frac{\partial C}{\partial W_{1,2}^{(2)}} \tag{9}$$

$$b_1^{(1)} \to b_1^{(1)} - h\frac{\partial C}{\partial b_1^{(1)}} \tag{10}$$

$$b_2^{(1)} \to b_2^{(1)} - h\frac{\partial C}{\partial b_2^{(1)}} \tag{11}$$

$$b_1^{(2)} \to b_1^{(2)} - h\frac{\partial C}{\partial b_1^{(2)}} \tag{12}$$

$$\frac{\partial C}{\partial W_{1,1}^{(1)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \cdot W_{1,1}^{(2)}\sigma'(z_1^{(1)})\, x_i \tag{13}$$

$$\frac{\partial C}{\partial W_{2,1}^{(1)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \cdot W_{1,2}^{(2)}\sigma'(z_2^{(1)})\, x_i \tag{14}$$

$$\frac{\partial C}{\partial W_{1,1}^{(2)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \cdot \sigma(z_1^{(1)}) \tag{15}$$

$$\frac{\partial C}{\partial W_{1,2}^{(2)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \cdot \sigma(z_2^{(1)}) \tag{16}$$

$$\frac{\partial C}{\partial b_1^{(1)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \cdot W_{1,1}^{(2)}\sigma'(z_1^{(1)}) \tag{17}$$

$$\frac{\partial C}{\partial b_2^{(1)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \cdot W_{1,2}^{(2)}\sigma'(z_2^{(1)}) \tag{18}$$

$$\frac{\partial C}{\partial b_1^{(2)}} = \sum_{i=1}^{n}(y_i - f(x_i))\sigma'(z_1^{(2)}) \tag{19}$$