The Mathematics

The mathematics of forward propagation is as follows: Let the weight between neuron $i$ and neuron $j$ be $w_{ij}$. Let $a_i$ be the output of neuron $i$ (the bias term is generated via neuron $0^*$) and $\text{net}_j$ be the sum of the inputs to neuron $j$.

$$\text{net}_j = \sum_{i=0}^{n} w_{ij}a_i$$

where $n$ is the number of neurons feeding into neuron $j$.

$$a_j = f(\text{net}_j)$$

Backpropagation modifies the weights and biases within the network based on the error computed as the difference between the network output and the desired output. The cost of the first neuron of the dual network is:

$$e_i = t_i - a_i$$

$$\delta_i = e_i f'(\text{net}_i)$$

The output of the neuron $i$ in the next layer (hidden layer) $\delta_i$ is computed as follows ($\delta_j$ is from the previous layer):

$$\delta_i = f'(\text{net}_i) \sum_j \delta_j w_{ij}$$

Finally, weight modification:

$$\Delta w_{ij} = w_{ij} + \eta \delta_j a_i$$

where $\eta$ is the learning rate.