1 Finite Differences

As the ϵ decrease in Finite Difference approach result is derivative of cost function.

$$C'(w) = \lim_{\epsilon \to 0} \frac{C(w+\epsilon) - C(w)}{\epsilon} \tag{1}$$

2 Cost function

Cost function is used to calculate the output of the network and using the output we can calculate how accurate the function is?
Cost function for a single layer can be written as:

2.1 Cost Function for single neuron

$$C(w) = \frac{1}{n} \sum_{i=1}^{n} (x_i w - y_i)^2 \tag{2}$$

$$C'(w) = \frac{1}{n} \sum_{i=1}^{n} 2x \left(k_i w - y_i\right) \tag{3}$$