**Summary of Backpropagation**

Notation:

* : layer number
* : final layer
* : output neuron number
* : input neuron number
* : weight vector
* : bias vector
* : activation function
* : cost function
* : Hadamard product
* : number of training examples in mini-batch
* : learning rate (
* : set of training examples
* : desired output

Algorithm:

1. **Input** **:** Set the corresponding activation for the input layer.
2. **Feedforward:** For each , compute and .
3. **Output error :** Compute the error vector . If the cost function is the mean squared error, then .
4. **Backpropagate the error:** For each , compute .
5. **Output:** For each , and . If using stochastic gradient descent, then and . Note that