# Feedforward equations

Where n = number of nodes in layer (L)

p = number of nodes in previous layer (L-1)

# Backpropagation with gradient descent

## calculate dj\_Dw and dj\_db for output layer

## calculate dj\_Dw and dj\_db for the rest of the layers

## UPDATE WEIGHTS AND BIASES OF EACH LAYER

Where is the learning rate

## IN TERMS OF MATRICIES

### OUTPUT LAYER

### REST OF THE LAYERS

Old

New

Repeat this back to first layer while setting the new to be used to avoid recalculations.

## Update weights and biases

For each layer:

Similar statements can be said for the biases