1.2 BACKPROPAGATION IN FEEDFORWARD MULTI-LAYER NETWORKS

Consider the neural network at Figure 1.1. Each hidden neuron uses a sigmoid activation function $\sigma(t) = 1/\left(1+\exp(-t)\right)$. The loss is the mean square error and the datapoints are $\left\{(x_1, x_2, y)_i\right\}_{i=1,2,3} = \{(1,1,-1),(0,-2,1),(-1,1,1)\}$. Input units are green, biases are yellow, hidden units blue and the output unit red. The weights can be found on the corresponding arrows. Compute the gradient through backpropagation analytically and don't forget to give the corresponding development. After one step of (simple, computed on the batch) gradient descent with learning rate 1, what are the new weights ? No need of a computer here, all can be done with pen and paper.

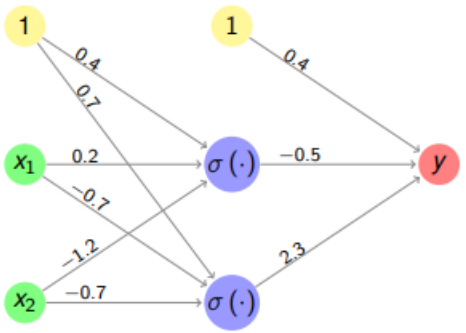

Figure 1.1: Neural network to perform the back-propagation on.

Hints. $\mathrm{d}(\sigma(t))/\mathrm{d}t = \sigma(t)(1-\sigma(t))$. Verifying that the new loss has indeed decreased is also a good hint.