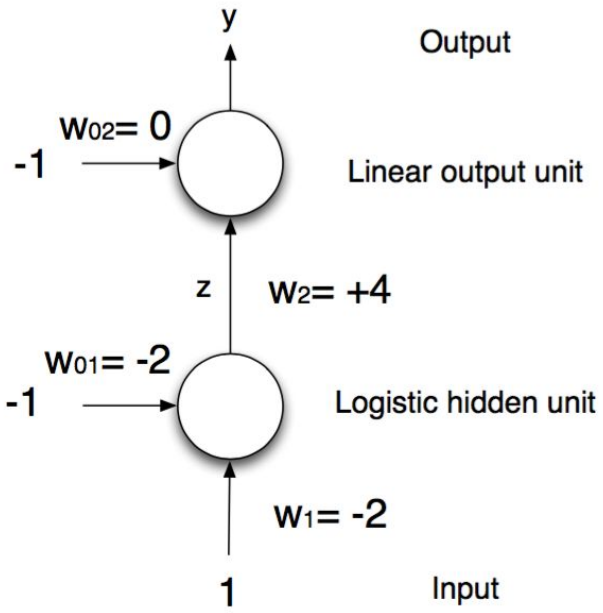
PROBLEM 23

14 Backpropagation

Here you see a very small neural network: it has one input unit, one hidden unit (logistic), and one output unit (linear).



Let's consider one training case. For that training case, the input value is 1 (as shown in the diagram), and the target output value $t = 1$. We're using the following loss function:

$$E = \frac{1}{2}(t - y)^2$$

Please supply numeric answers; the numbers in this question have been constructed in such a way that you don't need a calculator. Show your work in case of mis-calculation in earlier steps.

(a) What is the output of the hidden unit for this input?

(b) What is the output of the output unit for this input?

(c) What is the loss, for this training case?

(d) What is the derivative of the loss with respect to w_2, for this training case?

(e) What is the derivative of the loss with respect to w_1, for this training case?

(f) With sigmoidal activation, the derivative with respect to w_1 and w_2 are

$$\frac{\partial E}{\partial w_2} = -(t - y)z, \text{ and } \frac{\partial E}{\partial w_1} = -(t - y) \cdot w_2 \cdot z \cdot (1 - z) \cdot x.$$

Assume that we now use the rectified linear unit (ReLU) as our activation (or a *ramp* function). This means that $z = \max(0, w_1 x + w_{01})$. What is the derivative of the loss with respect to w_1 and w_2 *at differentiable points* with ReLU? Don't use numerical value for this question.