Chapter 6 lecture questions

Q1: Show that

$$\tanh(x) = 2\ \text{logistic}(2x) - 1\,. \qquad (1)$$

Solution:

$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \left(\frac{e^{-x}}{e^{-x}}\right) = \frac{1 - e^{-2x}}{1 + e^{-2x}} = \frac{2 - (1 + e^{-2x})}{1 + e^{-2x}} = 2\left(\frac{1}{1 + e^{-2x}}\right) - 1 = 2\ \text{logistic}(2x) - 1$

Q2: Given 50 observations in time, you want to use a 1-hidden layer MLP to model y as a function of $\mathbf{x}$, where $\mathbf{x}$ contains 6 predictors. How many hidden neurons can you use without violating $N_\mathrm{p} < N$?

Answer: Number of hidden neurons allowed is 1,2,3,4,5,6.

Solution:

For a m-m_2-1 MLP network, the number of parameters from lecture slide is $N_\mathrm{p} = (m+1)m_2 + (m_2+1)$. Here $m = 6$, so $N_\mathrm{p} = 7m_2 + (m_2+1) = 8m_2 + 1$. Try a range of m_2 values, and find the corresponding N_p values:

$m_2 =$	1	2	3	4	5	6	7
$N_p =$	9	17	25	33	41	49	57

To keep $N_\mathrm{p} < N = 50$, the highest m_2 allowed is 6. So the number of hidden neurons allowed is 1,2,3,4,5,6.

Q3: Given three outputs, $a_1 = -1$, $a_2 = 1$ and $a_3 = 5$. What does

$$y_k = \frac{\exp(a_k)}{\sum_{k'} \exp(a_{k'})}$$

give for the values y_1, y_2 and y_3?

Note: Matlab provides a function `softmax`, see

`http://www.mathworks.com/help/nnet/ref/softmax.html`

Answer: $y_1 = 0.0024$, $y_2 = 0.0179$ and $y_3 = 0.9796$.

Solution:

In Matlab:

```
>> a = [-1; 1; 5];     % Note semi-colons, not commas, in the brackets.
>> y = softmax(a)
y =
0.0024
0.0179
0.9796
```

Indeed, although a_3 is not that much bigger than the other two numbers, the softmax function gave $y_3 \approx 1$ and the other two $y_k \approx 0$.