Pattern Recognition - Fall 2024

Homework 3
Due date: Tuesday, December 31, 2024

1. (40 %) Consider a three-layer backpropagation net with d input units, n_H hidden units, c output units, and bias. Please derive the updating rules by using the logistic function

$$f(x) = \frac{1}{1+e^{-x}}$$

as your activation function and the cross entropy function as your criterion function.

   (a) Show that the updating rule for the hidden-to-output weights is as follows.

$$w_{kj} \leftarrow w_{kj} - \eta \frac{\partial J}{\partial w_{kj}} = w_{kj} + \eta(t_k - z_k)y_j$$

   (b) Show that the updating rule for the input-to-hidden weights is as follows.

$$w_{ji} \leftarrow w_{ji} - \eta \frac{\partial J}{\partial w_{ji}} = w_{ji} + \eta \sum_{k=1}^{c} (t_k - z_k) w_{kj} \cdot y_j(1-y_j) \cdot x_i$$

2. (60 %) **Computer Exercise 4**
Write a MATLAB function which trains a 2-2-1 network by means of batch back-propagation to solve the XOR problem. The syntax of this function is as follows:

```
[z W1 W2] = forward_backward(W1, W2, eta)
```

where `z` is a 1×4 matrix representing the outputs generated by feeding 4 input samples to the network, `W1` is a 2×3 matrix representing the input-to-hidden weights, `W2` is a 1×3 matrix representing the hidden-to-output weights and `eta` is the learning rate. Please use the updating rules derived in the previous question.

   (a) Determine the output matrices `[z W1 W2]` from the function `forward_backward(W1, W2, eta)` when all weights are set to 0.5 and the learning rate is set to 0.1.

   (b) Initialize all weights randomly in the range $-1 \leq w \leq 1$ and set the learning rate to 0.1. Plot the learning curve (i.e., the following function as a function of epoch) for 100 epochs.

$$J = \sum_{m=1}^{4} \{-t^m \ln(z^m) - (1-t^m)\ln(1-z^m)\}$$

   Note that t^m and z^m denote the correct output and the network's output of the m-th sample, respectively.