Assignment 1

1 Gradient Descent

Consider the function

$$f(x) = 1.5x_1^2 + x_2^2 - 2x_1x_2 + 2x_1^3 + 0.5x_1^4$$

Write a program to find the minimum of the function using gradient descent, without using a library-provided function of gradient descent. In other words, implement gradient descent. Start at $x_1=1, x_2=-3$ and follow the direction of the negative gradient. The output should be a list of 5 values per line: $x_1 \ x_2 \ f(x_1,x_2) \ \frac{\partial}{\partial x_1}(x_1,x_2) \ \frac{\partial}{\partial x_2}(x_1,x_2)$

For example:

. . .

Please answer the following questions:

- a) What is the actual minimum?
- b) When does the algorithm stop?
- c) How did you set the step size?

2 Neural Network Training

Consider the function found in the file on the course homepage http://deeplearningcourse.onai.com

Train a neural network with one input layer, one hidden layer with 5 nodes, and one output layer with 2 nodes. The activation function for the hidden layer shall be the logistic sigmoid

$$h(x) = \frac{2}{1 + e^{-x^2}} - 1$$

. The output layer shall be the softmax normalization with a cross entropy loss function.

The output of your script should be one line for each training sample indicating the sample, the network output, and the error function, such as

$$x_1$$
 x_2 $o_1(x)$ $o_2(x)$ $err(x)$

Please answer the following questions:

- a) When does the algorithm stop?
- b) How did you set the step size?

Submission Notes

This assignment may be completed as a group (maximum 3 students per group). All students in a group will receive the same score. Please submit a printout of your answers to Task 1 and 2 along with a printout of your source code and the output of your programs (not more than 2 pages—please only show the beginning and the end of the output). Your submission is due at the start of class, Feb. 14, 2018.