<u>Home Work Project</u>

Part 1. *<u>Realization of BackProp for identification of handwritten digits.</u>*

On moodle find a Matlab file with 5000 examples of handwritten digits. The examples were taken from MNIST archive: http://yann.lecun.com/ exdb/mnist

The file has two variables : X has 5000 columns, each column is a digit that is drawn on a grid of 20 by 20 pixels. Y: is the digit itself.

<u>A.</u> Draw a random digit from the file and verify that indeed it matches the data at Y. In order to draw the digits you should normalize the values according to the maximal value in the columns. In matlab you can use colormap(grey) and images, a suitable range for the black and white is [-1,1]. Please submit the outcome.

<u>B.</u> For the Neural network you should supply 400 entries + Bias. 25 neurons at the middle layer + Bias. At the final Layer should be 10 neurons, one for each possible number. Build the network and use BackProp for the computation of the minima of the potential/cost function. Please note that your function should return two values **(i)** the value of the potential for a specific set of parameters and **(ii)** vector of the potential derivatives. Remember to reshape the data from network topology to linear vectors.

<u>B.1</u> Perform a single run of your implementation of BackProp and compare it to a direct numerical computation of the network potential. Make sure that the obtained values are close and plot the differences for the two methods. Remember to randomly initialize the network. fminunc can be unstable for this problem, it is sometimes better to use fuming. This function has MaxIter as an option (adjustable at the function code), plot the value of potential as you change this parameter from 10 to 500.

<u>C.</u> Include a regularization part to your potential and repeat B.

<u>D.</u> 1. For growing values of λ (in regularization) train the neural network (use high MaxIter) . Plot for each λ the resulting value of the potential as obtained for Cross Validation Data (20% of total Data). Is there an optimal value?

<u>D.2.</u> Use the last 20% of the data for an optimal λ value and state the value of the potential of the trained network and what is the success rate.

Part 2. *Anomalous Diffusion and Rare Event Algorithm*

A. Generate N uniformly distributed random numbers on (0,1) and use the transformation method $y = cx^{-\alpha}$ in order to obtain power-law distributed random variables. **(i)** Choose α such that the decay of the distribution $P(y) \sim y^{-\frac{3}{2}}$ when $y \to \infty$. **(ii)** Plot on the log-log scale the distribution $P(y)$ as obtained from the simulated data, for $N = 10^2, 10^3, 10^4, 10^5, 10^6, 10^7$. **(iii)** By using the Maximal-likelihood method measure the values of the power-law exponent and plot it as a function of N. Also plot the estimated error for the exponent as a function of N.

B. On a one-dimensional lattice a random-walker is performing a unit step +1 to the right with probability $p < 1/2$ and a unit step to the left -1 with a probability $1 - p$, if the position of the random-walker (X) is positive, i.e $X > 0$. If $X < 0$, the probability for a step to the right is $1 - p$ and p is the probability for a step to the left. If $X = 0$ the probabilities are symmetric, half and half. **(i)** Simulate the process and plot the first-passage time distribution to reach $X = 20$ as a function of n, steps performed by the random walker. **(ii)** Plot this distribution (make it smooth) for $p = 0.45, 0.4, 0.3, 0.2$.

Plot the number of realizations (i.e. different random-walkers) you used for each p.

C. Use the STEPS rare events algorithm for the simulations in B. **(i)** Plot the results for different values of Q, the biasing part of the algorithm. What Q should you use? **(ii)** Modify the STEPS algorithm such that the biasing is performed by a moving absorbing boundary, it starts at $X = -1$ and moves one step to the right at each step of the random walker. If one of the possible positions of the random walker equals, or to the left, to the position of the absorbing boundary, terminate this option. What is Q in STEPS algorithm for such criteria for choosing possible paths? Use this modified version of STEPS in order to plot the probability distribution of first-passage time to $X = 100$ as a function of n, only for $p = 0.2$. Make it smooth. How many realizations you had to use. **(iii)** Try to obtain the results of (ii) without the STEPS algorithm. Plot the outcome and explain it.