Homework #2

Important note: Only soft copy will be accepted. Please supply the code in your answer if computer experiment is involved.

Q1. Rosenbrock's Valley Problem

Consider the Rosenbrock's Valley function:

$$f(x,y) = (1-x)^2 + 100(y-x^2)^2$$

which has a global minimum at (x,y) = (1,1) where f(x,y) = 0. Now suppose the starting point is randomly initialized in the open interval (0, 1) for x & y, find the global minimum using:

a). Steepest (Gradient) descent method

$$w(k+1) = w(k) - \eta g(k)$$

with learning rate η = 0.001. Record the number of iterations when f(x,y) converges to (or very close to) 0 and plot out the trajectory of (x,y) in the 2-dimensional space. Also plot out the function value as it approaches the global minimum. What would happen if a larger learning rate, say η = 0.1, is used?

b). Newton's method

$$\Delta w(n) = -H^{-1}(n)g(n)$$

Record the number of iterations when f(x,y) converges to (or very close to) 0 and plot out the trajectory of (x,y) in the 2-dimensional space. Also plot out the function value as it approaches the global minimum.

Q2. Function Approximation

Consider using MLP to approximate the following function:

$$y = 1.2\sin(\pi x) - \cos(2.4\pi x) \quad \text{, for } x \in [-1,1].$$

The training set is constructed by dividing the range [-1, 1] using a uniform step length 0.05, while the test set is constructed by dividing the range [-1, 1] using a uniform step length 0.01. You may use a MLP and do the following experiments:

a). Use the sequential mode with BP algorithm and experiment with the following different structures of the MLP: 1-n-1 (where n = 1,2,...,10,50). For each architecture plot out the outputs of the MLP for the test samples after training and compare them to the desired outputs. Try to determine whether it is under-fitting, proper fitting or over-fitting. Identify the minimal number of hidden neurons from the experiments, and check if the result is consistent with the guideline given in the lecture slides. Compute the outputs of the MLP when x=-1.5 and 1.5, and see if the MLP can make reasonable predictions outside of the range of the input limited by the training set.

b). Use the batch mode with regularization to repeat the above procedure.