ELE785- NEURAL NETWORKS **14.04.2024**

Homework I

Q1. Minimum Norm. Find least square estimator $\boldsymbol{\theta}_{ls}$ for m<n (i.e. there are more variables than equations since **m** is the number of data pairs recorded and **n** is the number of parameters) under the minimum norm consideration by stating a minimization problem as follows:

minimize $|\boldsymbol{\theta}|$
subject to: $A\boldsymbol{\theta}= y$ (with variable $\theta \in R^n$)

Q2. RLS Study. a**)** Find the least square estimator $\boldsymbol{\theta} =[\boldsymbol{\theta}_0\ \boldsymbol{\theta}_1\ \boldsymbol{\theta}_2]$ for the model
$y(t)= \boldsymbol{\theta}_0 + \boldsymbol{\theta}_1 x_1+ \boldsymbol{\theta}_2 x_2+ e'(t)$ proposed for the data set given in the Table.1 below,

b) Use LMS algorithm to find $\boldsymbol{\theta}=[\ \boldsymbol{\theta}_0\ \boldsymbol{\theta}_1\ \boldsymbol{\theta}_2\ ]$ for the same model. Choose the learning rate carefully for the convergent iteration. Plot each variable $\boldsymbol{\theta}_i$ during the adaptation. Compare your results with the Wiener's Optimal Solution ($\boldsymbol{\theta}^*=\mathbf{R}^{-1}\ \mathbf{p}$).

x1	1	2	2	2	3	3	4	5	5	5	6	7	8	8	9
x2	2	5	3	2	4	5	6	5	6	7	8	6	4	9	8
y	2	1	2	2	1	3	2	3	4	3	4	2	4	3	4

Table.1

Q3. Derivative Based Optimization. Consider minimization problem for the objective function,

$$E_1(\theta)=E(x,y)= \boldsymbol{\theta}_1^2+4\boldsymbol{\theta}_2^{\ 2}+2\boldsymbol{\theta}_1\boldsymbol{\theta}_2+ \boldsymbol{\theta}_1-11\boldsymbol{\theta}_2 \qquad \theta=[\boldsymbol{\theta}\mathbf{1},\ \boldsymbol{\theta}\mathbf{2}]$$

a) Apply the following four descent methods to minimize these quadratic functions, using optimal step size η^* (line minimization) when necessary and initiate those algorithms for the initial condition as $\theta(0) =[1,1]$

- The steepest descent method
- Newton's method
- DFP quasi-Newton method
- Fletcher-Reeves's conjugate gradient method

Comment on the results. (Approximate $E(\theta)$ by using Taylor's theorem: $E(\theta)= 0.5\ \theta^T H\theta+g^T\theta+c$)

b) Use fixed η values (η=0.1, η=0.3, η=0.5) for the steepest descent method. Give your comments on the results.

Note: Plot each step of iterations on the contour surface in the range of $\boldsymbol{\theta}_1 \in [-2,10]$, $\boldsymbol{\theta}_2 \in [-4, 2]$.