ECE750/SDE625

MIDTERM EXAM

Due electronically (online upload) on June 29, 2012 (5pm)

Instructions:

Use this doc file template when producing the midterm answers. Upload under Midterm folder (in the course ftp site at: ftp://pami.uwaterloo.ca with username:ece750s10; password:750s10*) the following files (userid.doc, userid.pdf, userid.zip). The zip file should have all files (doc, pdf and commented Matlab files). This work should be done completely on a strictly individual basis.*

Problem 1

Part a)

Consider an *n*-ary fuzzy relation $R(x, y, z, \ldots)$, which has *n* independent variables (coordinate axes), $x, y, z, \ldots$. How many different projections are possible for this relation, in various subspaces? Explain how you arrived at this number.

Consider a discrete ternary ($n = 3$) fuzzy relation (rule base) given by the following "pages" of binary (x_i, y_j) relation along the z axis:

$$R_{z_1}(x_i, y_j) = \begin{bmatrix} 0.5 & 0.4 & 0.2 \\ 0.3 & 0.8 & 0.5 \\ 0.1 & 0.6 & 0.6 \end{bmatrix}; \; R_{z_2}(x_i, y_j) = \begin{bmatrix} 0.6 & 0.5 & 0.3 \\ 0.4 & 1.0 & 0.6 \\ 0.2 & 0.6 & 0.8 \end{bmatrix};$$

$$R_{z_3}(x_i, y_j) = \begin{bmatrix} 0.4 & 0.3 & 0.1 \\ 0.2 & 0.7 & 0.3 \\ 0.0 & 0.5 & 0.5 \end{bmatrix}$$

Determine all possible projections of this relation

Part b)

Show that *max*[0, *x*+ *y* - 1] is a t-norm. Also, determine the corresponding t-conorm (i.e., s-norm).

Problem 2

Part a)

Show that the optimal learning rate η_{opt} for minimizing the quadratic error function $E(w(k)) = 0.5 w^T Q w$ **is:**

$\eta_{opt} = (w^T Q^2 w / w^T Q^3 w)|_{(k)}$

(*Hint:* Express the error function at step $(k+1)$, and minimize this expression with respect to the learning rate η. Use the fact that for quadratic forms $Q = Q^T$)

Part b)

Using results in problem 1, perform two iterations of the optimal gradient algorithm to find the minimum of $E(\boldsymbol{w}) = 2w_1^2 + 2w_1 w_2 + 5w_2^2$, where $\boldsymbol{w} = [w_1 \quad w_2]^T$ is the weight vector of dimension 2. The starting point is $\boldsymbol{w}(0) = [2 \quad -2]^T$. Find $\boldsymbol{w}(1)$ and show that the gradients $\nabla E(\boldsymbol{w}(0))$ and $\nabla E(\boldsymbol{w}(1))$ are perpendicular.

Problem 3

We need to train a Radial Basis Function network for obtaining the output of the following two-to-one mapping function:

$$y(x_1, x_2) = \sin(2\pi x_1)\cos(0.5\pi x_2)$$

a) Set up two sets of data, each of which consists of 100 input-output patterns, one for network training and the other for testing (with which to compare performance with the true function). The input-output data are obtained by randomly varying the input variables (x_1, x_2) within the interval [-1,1].

b) Use two type RBF networks: one with Gaussian MF's and the other with exponential MF's (logistic). Analyse the results by changing the number of nodes from 10 to 100 (in steps of 10). Compare the outcome of both networks (with different basis functions). Compare the performance of the networks in terms of execution time and accuracy (mean squared errors.). Discuss your results

Problem 4

The dynamics of a time series is represented by the following nonlinear difference equation:

$$\begin{aligned} y(t) = {} & (0.3 - 0.6\exp(-y^2(t-1)))y(t-1) - \\ & (0.8 + 0.9\exp(-y^2(t-1)))y(t-2) + 0.3\sin(\pi y(t-1)) \end{aligned}$$

a) Considering y(t) as the output of the time series and $y(t-1)$ and $y(t-2)$ as its inputs with $(y(t-1), y(t-2)) \in [-1.8 \quad 1.8]^2$ display the surface output. Note the linear behaviour of $y(t)$ as a function of $y(t-2)$ when $y(t-1)$ is held at 0.

b) Design a feed forward neural network to estimate the output surface. You may wish to use 500 sample data for training and testing purposes (70% of which for training and 30% for testing). Use several scenarios to find the best network (varying the number of hidden layers from 1 to 3 and varying the number of nodes in each layer let say from 2 to 58 in steps of 8). Display your results in a table and discuss the results.

c) Compare the results obtained in (b) with those you obtain using a RBF network with 30, 60, 90 hidden nodes, respectively. Comparison is made in terms of Root Mean Square Error (RMSE).

Problem 5 (Linear Classifiers)

We need to develop a neurocomputing based classifier for three various but related products. The collected data are the results of a chemical analysis of liquid products grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 12 constituents found in each of the three types of products. The data file provided is in text format and has thirteen dimensions, the first of which determines the class of products (product '1', product '2', product '3'), which should serve as the output of the neurocomputing classifier. The remaining ones determine the input of the classifier and has 12 constituents as:

1)Malic acid
2) Ash
3) Alcalinity of ash
4) Magnesium
5) Total phenols
6) Flavanoids
7) Nonflavanoid phenols
8) Proanthocyanins
9)Color intensity
10)Hue
11)OD280/OD315 of diluted liquid
12)Proline

5.1 Design a classifier of your choice (neural network based or SVM based), vary its parameters and try to find the best possible classification performance (a table illustrating various results as parameters are varied would be preferred).

5.2 Once this is done, classify (determine to which product they belong) the following entries each of which has 12 attributes:

- **a)** 13.72; 1.43; 2.5; 16.7; 108; 3.4; 3.67; 0.19; 2.04; 6.8; 0.89; 2.87; 1285
- **b)** 12.04; 4.3; 2.38; 22; 80; 2.1; 1.75; 0.42; 1.35; 2.6; 0.79; 2.57; 580
- **c)** 14.13; 4.1; 2.74; 24.5; 96; 2.05; 0.76; 0.56; 1.35; 9.2; 0.61; 1.6; 560

Hint for implementation: You may wish to calibrate all input data to be all between 0 and one. Also from the set of data choose 75% of the data from product 1, product 2 and product 3 as training data and remaining 25% remaining as testing. You can use existing Matlab libraries or other libraries to create the classifier, which code needs to be appended to the solution. In case you make use of SOM (Self Organizing Map), please disregard the classes and assume three clusters.