Midsem Examination **AI and Neuro-Fuzzy Theory AT74.05** **March 11, 2016**

Time: 9:00-11:00 h. **Open Book**
Marks: 100

Attempt all questions.

Q.1 Consider the influence of different percentage of training in an ADALINE network which is used to determine the parameters; m and c, of a liner relation, $y = mx + c$, when x is the input and y is the output. The training set consists of the following data.

x	1	2	3	4
y	6.1	9.9	14.1	17.9

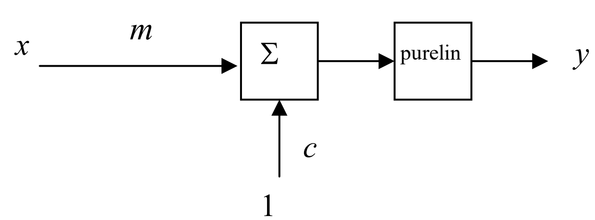
(a) Draw the ADALINE network that can be used to determine the parameters. (4)

(b) If all the data are presented equally, determine the parameters by LMS algorithm. (8)

(c) If the first data is presented 50%, the second data is presented 30%, the third data is presented 15%, and the last data is presented 5%, determine the parameters by LMS algorithm. (8)

Solution

(a)



(b)

$$F(x) = E[t^2] - 2x^T E[tz] + x^T E[zz^T]x = c - 2x^T h + x^T Rx \quad (1)$$

$$c = \frac{1}{4}\left(6.1^2 + 9.9^2 + 14.1^2 + 17.9^2\right) = 163.61 \quad (2)$$

$$h = \frac{1}{4}\left(6.1\begin{bmatrix}1\\1\end{bmatrix} + 9.9\begin{bmatrix}2\\1\end{bmatrix} + 14.1\begin{bmatrix}3\\1\end{bmatrix} + 17.9\begin{bmatrix}4\\1\end{bmatrix}\right) = \begin{bmatrix}34.95\\12.00\end{bmatrix} \quad (3)$$

$$R = \frac{1}{4}\left(\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}^T + \begin{bmatrix}2\\1\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}^T + \begin{bmatrix}3\\1\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix}^T + \begin{bmatrix}4\\1\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix}^T\right) = \begin{bmatrix}7.50 & 2.50\\2.50 & 1\end{bmatrix} \quad (4)$$

The minimum point is the stationary point of the quadratic function.

$$\begin{bmatrix} m \\ c \end{bmatrix} = R^{-1}h = \begin{bmatrix} 7.50 & 2.50 \\ 2.50 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 34.95 \\ 12.00 \end{bmatrix} = \begin{bmatrix} 3.96 \\ 2.10 \end{bmatrix} \quad \underline{(5)}$$

(c)

$$F(x) = E[t^2] - 2x^T E[tz] + x^T E[zz^T]x = c - 2x^T h + x^T Rx \quad (6)$$

$$c = \left(0.5\times 6.1^2 + 0.3\times 9.9^2 + 0.15\times 14.1^2 + 0.05\times 17.9^2\right) = 93.85 \quad (7)$$

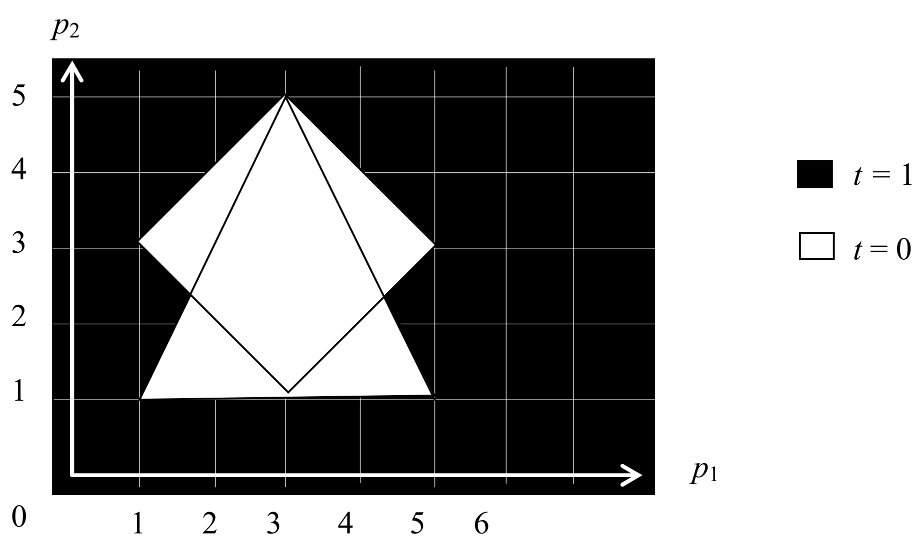
$$h = \left(0.5\times 6.1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0.3\times 9.9\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 0.15\times 14.1\begin{bmatrix} 3 \\ 1 \end{bmatrix} + 0.05\times 17.9\begin{bmatrix} 4 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 18.915 \\ 9.03 \end{bmatrix} \quad (8)$$

$$R = \left(0.5\times\begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T + 0.3\times\begin{bmatrix} 2 \\ 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}^T + 0.15\times\begin{bmatrix} 3 \\ 1 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix}^T + 0.05\times\begin{bmatrix} 4 \\ 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}^T\right) = \begin{bmatrix} 3.85 & 1.75 \\ 1.75 & 1 \end{bmatrix} \quad (9)$$

The minimum point is the stationary point of the quadratic function.

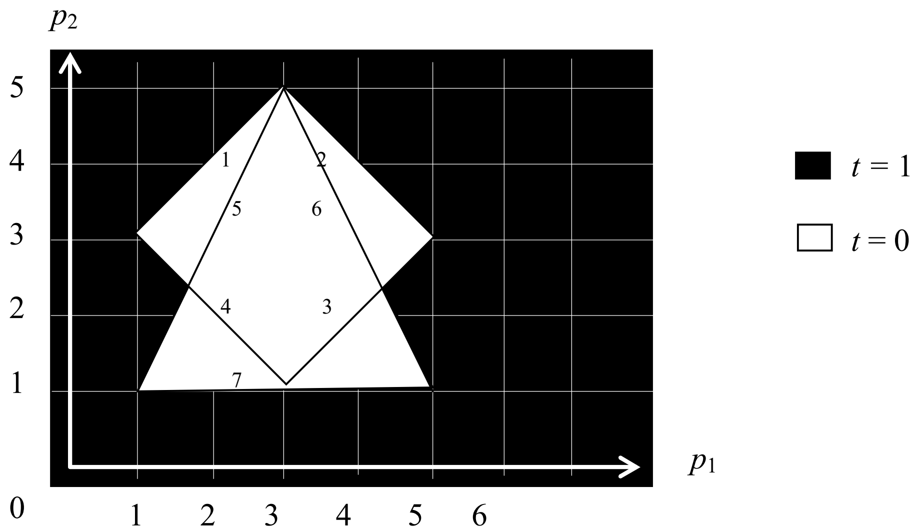
$$\begin{bmatrix} m \\ c \end{bmatrix} = R^{-1}h = \begin{bmatrix} 3.85 & 1.75 \\ 1.75 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 18.915 \\ 9.03 \end{bmatrix} = \begin{bmatrix} 3.95 \\ 2.11 \end{bmatrix} \quad \underline{(10)}$$

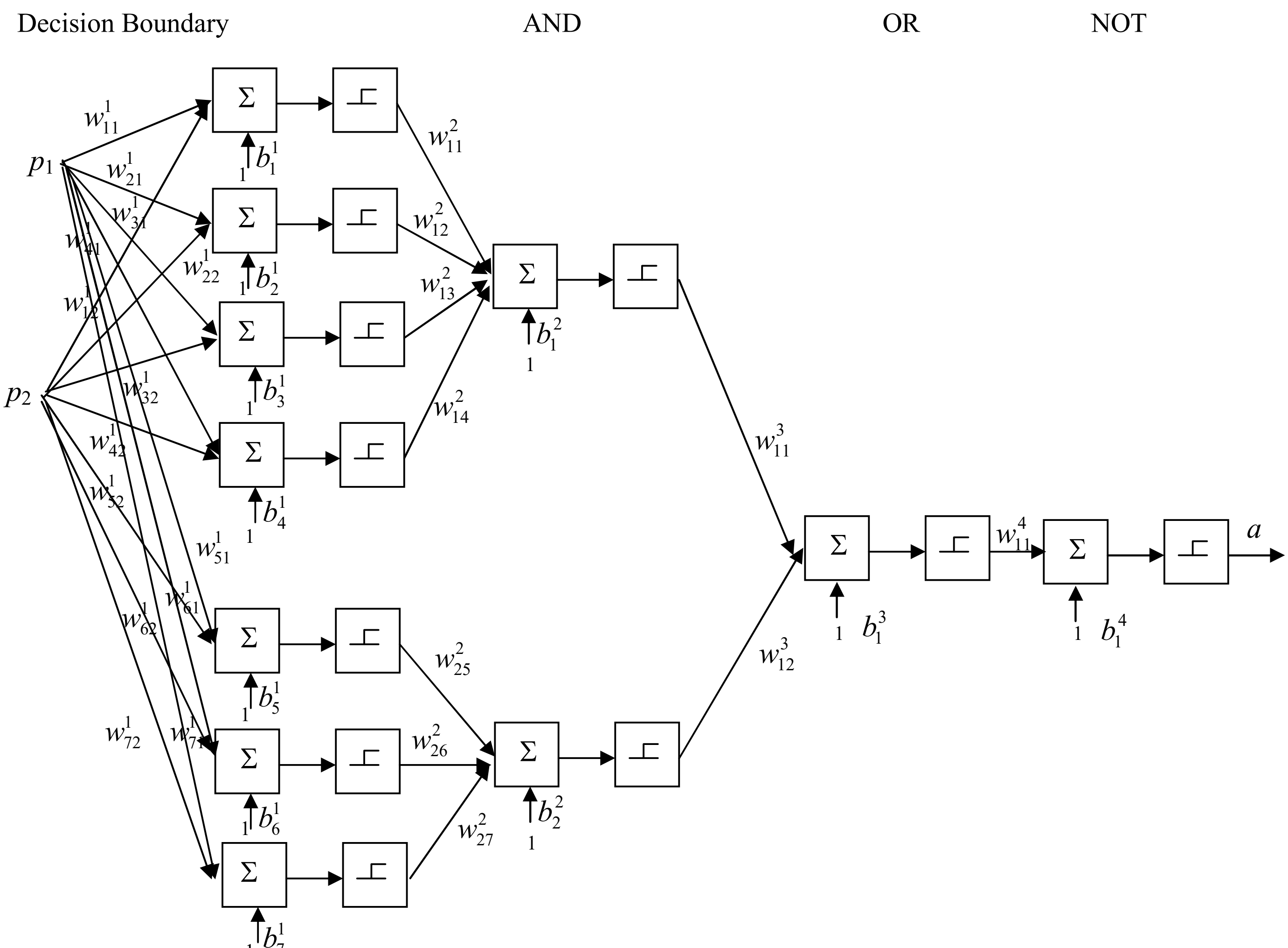
Q.2 Design a network which can generate the output according to the given training set correctly. Draw the selected network and determine all the required parameters. (20)



<u>Solution</u>

Multi-Layer Perceptron is selected. Each decision boundary is numbered as shown below.





Layer 1 is used to generate decision boundaries, layer 2 is used to AND the selected decision boundaries, layer 3 is used to OR the areas, layer 4 is used to negate the result.

For layer 1

Since **w** should point into the white area and should be perpendicular to decision boundary. Thus, select

For square-shape area,

$$\mathbf{w}_1^1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (1)$$

$$\mathbf{w}_2^1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad (2)$$

$$\mathbf{w}_3^1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (3)$$

$$\mathbf{w}_4^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (4)$$

For triangle-shape area,

$$\mathbf{w}_5^1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad (5)$$

$$\mathbf{w}_6^1 = \begin{bmatrix} -2 \\ -1 \end{bmatrix} \quad (6)$$

$$\mathbf{w}_7^1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (7)$$

b is determined from equating **n** = 0; and solve for **b**

For square-shape area,

$$n_1^1 = w_{11}^1 p_1 + w_{12}^1 p_2 + b_1^1 = 0 \quad (8)$$

at $p_1 = 3$, $p_2 = 5$;

$$1(3) - 1(5) + b_1^1 = 0; b_1^1 = 2 \quad (9)$$

$$n_2^1 = w_{21}^1 p_1 + w_{22}^1 p_2 + b_2^1 = 0 \quad (10)$$

at $p_1 = 3$, $p_2 = 5$;

$$-1(3) - 1(5) + b_2^1 = 0; b_2^1 = 8 \quad (11)$$

$$n_3^1 = w_{31}^1 p_1 + w_{32}^1 p_2 + b_3^1 = 0 \quad (12)$$

at $p_1 = 3$, $p_2 = 1$;

$$-1(3) + 1(1) + b_3^1 = 0; b_3^1 = 2 \quad (13)$$

$$n_4^1 = w_{41}^1 p_1 + w_{42}^1 p_2 + b_4^1 = 0 \quad (14)$$

at $p_1 = 3$, $p_2 = 1$;

$$1(3) + 1(1) + b_4^1 = 0; b_4^1 = -4 \quad (15)$$

For triangle-shape area,

$$n_5^1 = w_{51}^1 p_1 + w_{52}^1 p_2 + b_5^1 = 0 \quad (16)$$

at $p_1 = 3$, $p_2 = 5$;

$$2(3) - 1(5) + b_1^1 = 0; b_5^1 = -1 \quad (17)$$

$$n_6^1 = w_{61}^1 p_1 + w_{62}^1 p_2 + b_6^1 = 0 \quad (18)$$

at $p_1 = 3$, $p_2 = 5$;

$$-2(3) - 1(5) + b_6^1 = 0; b_6^1 = 11 \quad (19)$$

$$n_7^1 = w_{71}^1 p_1 + w_{72}^1 p_2 + b_7^1 = 0 \quad (20)$$

at $p_1 = 1$, $p_2 = 1$;

$$0(1) + 1(1) + b_7^1 = 0; b_7^1 = -1 \quad (21)$$

For layer 2

Since this is AND layer, if we select

$$w_{11}^2 = w_{12}^2 = w_{13}^2 = w_{14}^2 = 1 \quad (22)$$

$$w_{25}^2 = w_{26}^2 = w_{27}^2 = 1 \quad (23)$$

b must be selected such that only when all inputs are 1s, output is 1. Thus, select

$$b_1^2 = -3.5 \quad (24)$$

$$b_2^2 = -2.5 \quad (25)$$

For layer 3

Since this is OR layer, if we select

$$w_{11}^3 = w_{12}^3 = 1 \quad (26)$$

b must be selected such that even one input is 1, output is 1. Thus, select

$$b_1^3 = -0.5 \quad (27)$$

For layer 4

Since this is NOT layer, if we select

$$w_{11}^4 = -1 \tag{28}$$

b must be selected such that when the input is 1, output is 0. Thus, select

$$b_1^4 = 0.5 \tag{29}$$

The weights and biases which are not mentioned are set to zero.

Q.3 Pseudo-inverse rule for linear associator is applicable when the number of rows in the input vector is not less than the number of data in the training set.

(a) Is it possible to determine the weight matrix by pseudo-inverse rule when the training set consists of

$\{p_1 = [1], t_1 = [20]\}, \{p_2 = [2], t_2 = [30]\}, \{p_3 = [3], t_1 = [50]\}, \{p_4 = [4], t_1 = [80]\}$?

If yes, determine the weight matrix, then test by presenting all the inputs. If no, explain the reason. (10)

(b) If the number of input is extended to have at least equal number as training set by using non-linear operation of the input; for example using power operation as shown in the training set below.

$$\left\{p_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, t_1 = [20]\right\}, \left\{p_2 = \begin{bmatrix}2\\4\\8\\16\end{bmatrix}, t_2 = [30]\right\}, \left\{p_3 = \begin{bmatrix}3\\9\\27\\81\end{bmatrix}, t_1 = [50]\right\}, \left\{p_4 = \begin{bmatrix}4\\16\\64\\256\end{bmatrix}, t_1 = [80]\right\},$$

Is it possible to determine the weight matrix by pseudo-inverse rule or not? If yes, determine the weight matrix, then test by presenting all the inputs. If no, explain the reason. (10)

<u>Solution</u>

By Pseudoinverse rule,

$$\mathbf{W} = \mathbf{T}\mathbf{P}^+ \tag{1}$$

Where

$$\mathbf{P}^+ = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T \tag{2}$$

(a)

$$\mathbf{T} = \begin{bmatrix}20 & 30 & 50 & 80\end{bmatrix} \tag{3}$$

$$\mathbf{P} = \begin{bmatrix}1 & 2 & 3 & 4\end{bmatrix} \tag{4}$$

$$\left|\mathbf{P}^{\mathrm{T}}\mathbf{P}\right| = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \\ 4 & 8 & 12 & 16 \end{vmatrix} = 0 \tag{5}$$

Since matrix $\mathbf{P}^T\mathbf{P}$ is a singular matrix, there is no inverse of this matrix. Thus, pseudo-inverse matrix cannot be determined. It is not possible to determine the weight matrix by pseudo-inverse rule.

(a)

$$\mathbf{T} = \begin{bmatrix} 20 & 30 & 50 & 80 \end{bmatrix} \tag{6}$$

$$\mathbf{P} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \\ 1 & 16 & 81 & 256 \end{bmatrix} \tag{7}$$

$$\mathbf{W} = \mathbf{T}\left(\mathbf{P}^{\mathrm{T}}\mathbf{P}\right)^{-1}\mathbf{P}^{\mathrm{T}} = \begin{bmatrix} 36.67 & -24.17 & 8.33 & -0.83 \end{bmatrix} \tag{8}$$

$$\mathbf{W}p_1 = \begin{bmatrix} 36.67 & -24.17 & 8.33 & -0.83 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 20 \end{bmatrix} \tag{9}$$

$$\mathbf{W}p_2 = \begin{bmatrix} 36.67 & -24.17 & 8.33 & -0.83 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 8 \\ 16 \end{bmatrix} = \begin{bmatrix} 30 \end{bmatrix} \tag{10}$$

$$\mathbf{W}p_3 = \begin{bmatrix} 36.67 & -24.17 & 8.33 & -0.83 \end{bmatrix} \begin{bmatrix} 3 \\ 9 \\ 27 \\ 81 \end{bmatrix} = \begin{bmatrix} 50 \end{bmatrix} \tag{11}$$

$$\mathbf{W}p_4 = \begin{bmatrix} 36.67 & -24.17 & 8.33 & -0.83 \end{bmatrix} \begin{bmatrix} 4 \\ 16 \\ 64 \\ 256 \end{bmatrix} = \begin{bmatrix} 80 \end{bmatrix} \tag{12}$$

Q.4 Neural network can be applied for system identification of a dynamic system as explained by $\dot{x} = ax + bu$, when a and b are unknown parameters. The discrete version of the relation can be expressed by $x(k+1) = e^{a\tau}x(k) + bu(k)$ when τ is the sampling period and assumed 0.1 second.

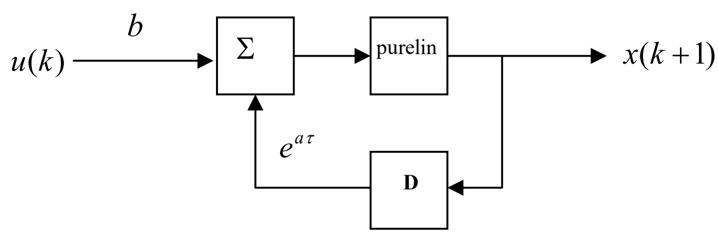
(a) Based on the discrete relation, design a neural network that can represented the relation. (5)

(b) If data is collected as shown in the below table, determine the parameters a and b. (15)

$u(k)$	2	5	-3	
$x(k)$	0	4	13.62	6.32

Solution

(a)



(b)

$$F(x) = E[t^2] - 2x^T E[tz] + x^T E[zz^T]x = c - 2x^T h + x^T Rx \quad (1)$$

$$c = \frac{1}{3}\left(4^2 + 13.62^2 + 6.32^2\right) = 80.48 \quad (2)$$

$$h = \frac{1}{3}\left(4\begin{bmatrix}2\\0\end{bmatrix} + 13.62\begin{bmatrix}5\\4\end{bmatrix} + 6.32\begin{bmatrix}-3\\13.62\end{bmatrix}\right) = \begin{bmatrix}19.05\\46.85\end{bmatrix} \quad (3)$$

$$R = \frac{1}{3}\left(\begin{bmatrix}2\\0\end{bmatrix}\begin{bmatrix}2\\0\end{bmatrix}^T + \begin{bmatrix}5\\4\end{bmatrix}\begin{bmatrix}5\\4\end{bmatrix}^T + \begin{bmatrix}-3\\13.62\end{bmatrix}\begin{bmatrix}-3\\13.62\end{bmatrix}^T\right) = \begin{bmatrix}12.67 & -6.95\\-6.95 & 67.17\end{bmatrix} \quad (4)$$

The minimum point is the stationary point of the quadratic function.

$$\begin{bmatrix}b\\e^{-0.1a}\end{bmatrix} = R^{-1}h = \begin{bmatrix}12.67 & -6.95\\-6.95 & 67.17\end{bmatrix}^{-1}\begin{bmatrix}19.05\\46.85\end{bmatrix} = \begin{bmatrix}2\\0.90\end{bmatrix} \quad (5)$$

$$\begin{bmatrix}b\\a\end{bmatrix} = \begin{bmatrix}2\\1\end{bmatrix} \quad (6)$$

Q.5 Discuss how to apply neural network to your research work. What kind of network is appropriate and how to prepare the training set? (20)