2、 (a) $g(x) = w^T x + b$, where $x = x_p + r\frac{w}{\|w\|}$, x_p is the projection of x at the hyperplane , r is the distance from x to the hyperplane.

$$\Rightarrow g(x) = w^T x_p + b + r\frac{w^T w}{\|w\|} = g(x_p) + r\|w\| = r\|w\|$$

$$\Rightarrow r = \frac{g(x)}{\|w\|}$$ To maximize the margin is equal to minimize the weight $\|w\|$.

$\Rightarrow$ primal optimization problem :

$$\min. \ J = \frac{1}{2} w^T w$$
$$\text{s.t.} \ d(i) \cdot [w^T x(i) + b] \geq 1$$

(b) Based on Lagrange multiplier method:

$$J(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha(i) [d(i)(w^T x(i) + b) - 1]$$

$$\frac{\partial J(w, b, \alpha)}{\partial w} = w - \sum_{i=1}^{N} \alpha(i) d(i) x(i) = 0 \Rightarrow w = \sum_{i=1}^{N} \alpha(i) d(i) x(i)$$

$$\frac{\partial J(w, b, \alpha)}{\partial b} = -\sum_{i=1}^{N} \alpha(i) d(i) = 0$$

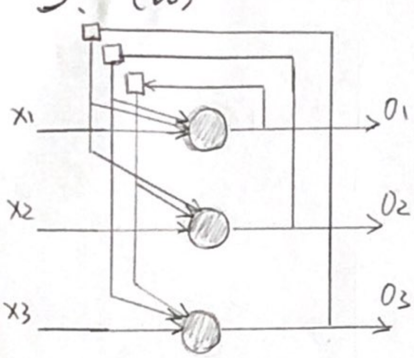
$$w^T w = \sum_{i=1}^{N} \alpha(i) d(i) x(i) \sum_{j=1}^{N} \alpha(j) d(j) x(j) = \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha(i)\alpha(j) d(i) d(j) x^T(i) x(j)$$

$$\Rightarrow J(w, b, \alpha) = \frac{1}{2} w^T w - \sum_{i=1}^{N} \alpha(i) d(i) w^T x(i) - \sum_{i=1}^{N} \alpha(i) d(i) b + \sum_{i=1}^{N} \alpha(i)$$
$$= \sum_{i=1}^{N} \alpha(i) - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha(i)\alpha(j) d(i) d(j) x(i) x(j) = Q(\alpha)$$

$\Rightarrow$ dual problem:

$$\min. \ Q(\alpha)$$
$$\text{s.t.} \ \sum_{i=1}^{N} \alpha(i) d(i) = 0$$
$$\alpha(i) \geq 0$$

3. (a)



$P_1 = [1 \ -1 \ 1 \ 1]^T$

$P_2 = [-1 \ 1 \ 1 \ 1]^T$

$P_3 = [1 \ 1 \ 1 \ 1]^T$

For the Hopfield neural network, the weight is calculated as:

$$W = \frac{1}{d}\sum_{i=1}^{N} S_i S_i^T - \frac{N}{d} I \quad , \ N=3, \ d=4$$

$$\Rightarrow \quad P_1 \cdot P_1^T = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} [1 \ -1 \ 1 \ 1] = \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}$$

$$P_2 \cdot P_2^T = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [-1 \ 1 \ 1 \ 1] = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

$$P_3 \cdot P_3^T = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [1 \ 1 \ 1 \ 1] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$W = \frac{1}{4}\begin{bmatrix} 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$$

, assume input is $P_1 = [1 \ -1 \ 1 \ 1]^T$ and a bi-polar activation function is used:

$$O(1) = \varphi(WX(1)) = \varphi\left(\frac{1}{4}[3 \ 1 \ 3 \ 3]^T\right) = [1 \ 1 \ 1 \ 1]^T$$

$$O(2) = \varphi(WO(1)) = \varphi\left(\frac{1}{4}[1 \ 1 \ 5 \ 5]^T\right) = [1 \ 1 \ 1 \ 1]^T = O(1)$$

$\therefore$ x converge to P_3 rather than P_1.

consider another input $P_2 = [-1 \ 1 \ 1 \ 1]^T$:

$$O(1) = \varphi(WX(2)) = \varphi\left(\frac{1}{4}[1 \ 3 \ 3 \ 3]^T\right) = [1 \ 1 \ 1 \ 1]^T = P_3$$

$\therefore$ x converge to P_3 rather than P_2.

The energy is $E = -\frac{1}{2} o^T W o$, for P_1:

$$E_1 = -\frac{1}{2} \times \frac{1}{4} \times [1 \ -1 \ 1 \ 1] \begin{bmatrix} 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = -1$$

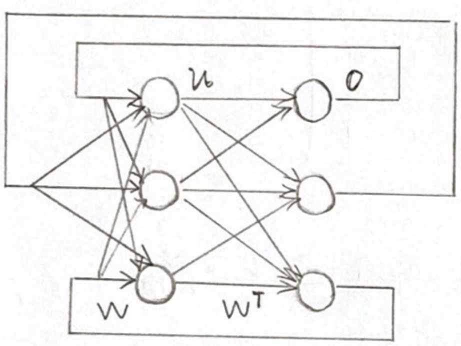
$E = -\frac{1}{2} o^T W o$, for $P_2 = [-1 \quad 1 \quad 1 \quad 1]^T$:

$$E_2 = -\frac{1}{2} \times \frac{1}{4} \times [-1 \quad 1 \quad 1 \quad 1] \begin{bmatrix} 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = -1$$

$$E_3 = -\frac{1}{2} \times \frac{1}{4} \times [1 \quad 1 \quad 1 \quad 1] \begin{bmatrix} 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = -\frac{3}{2}$$

The P_3 has lower energy, therefore the network tends to converge to P_3 rather than P_1 and P_2 because of the fact that low-energy state is always more stable.

(b)



For the BAM network, $W = \sum_{i=1}^{N} x_i y_i^T$

$$W_1 = P_1 P_2^T = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} [-1 \quad 1 \quad 1 \quad 1]$$

$$= \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix}$$

NANYANG TECHNOLOGICAL UNIVERSITY

SEMESTER 1 EXAMINATION 2014 - 2015

EE7207 – NEURAL AND FUZZY SYSTEMS

November/December 2014 Time Allowed: 3 hours

INSTRUCTIONS

1. This paper contains 5 questions and comprises 5 pages.
2. Answer ALL questions.
3. All questions carry equal marks.
4. This is a closed-book examination.

Unless specifically stated, all symbols have their usual meanings.

---

1. The performance of a radial basis function neural network is determined by a number of factors including the location of centre vectors of hidden layer neurons in the input space.

(a) Centre vectors are often selected based on the principle of input space coverage. Explain the rationale of this principle.

(5 Marks)

(b) Self-organizing map (SOM) neural networks can be employed to select centre vectors for RBF neural networks. Explain the rationale and describe the procedure.

(10 Marks)

(c) The number of neurons in the hidden layer is also an important consideration when using neural networks to solve pattern classification and regression problems. Discuss the implications of using unsuitable number of hidden layer neurons, and suggest a way to determine the suitable number of neurons.

bottom-up, forward selection

(5 Marks)

top-down, backward elimination

2. Support vector machine (SVM) is recognized as an effective learning machine for pattern classification.

(a) Taking binary linear separable patterns as an example, briefly describe the objective of SVM and the associated primal optimization problem.

(6 Marks)

(b) Derive the dual optimization problem of the primal optimization problem in Question 2(a).

(10 Marks)

(c) Discuss the differences in the design of Gaussian kernel SVMs and Gaussian radial basis function (RBF) neural networks.

(4 Marks)

3. There are three patterns:

$$P_1 = [1 \quad -1 \quad 1 \quad 1]^T$$

$$P_2 = [-1 \quad 1 \quad 1 \quad 1]^T$$

$$P_3 = [1 \quad 1 \quad 1 \quad 1]^T$$

(a) Design a Hopfield neural network to store the three patterns, and check whether the designed Hopfield neural network could correctly retrieve the three patterns. Explain the results from the energy point of view.

(9 Marks)

(b) Design a bi-directional memory (BAM) neural network to store the three patterns and check whether BAM neural network designed could correctly retrieve the three patterns.

(7 Marks)

(c) From load parameter point of view, give some suggestions to design reliable Hopfield neural network.

(4 Marks)

The load parameter is defined as $\alpha = \frac{N}{d}$, where N is the number of fundamental memories, d is the number of neurons in the network. The storage ability of the Hopfield neural network deteriorate with the increase of α, therefore, N should be small and d should be large.

4. (a) A t-conorm S is a binary operation on the unit interval that satisfies the following axioms for all $a, b, c \in [0, 1]$.

    (i) $S(0,a)=a$;

    (ii) $b \leq c$ implies that $S(a,b) \leq S(a,c)$

    (iii) $S(a,b) = S(b,a)$

    (iv) $S(a,S(b,c))=S(S(a,b),c)$

    Prove that the standard union $S(a, b) = max(a, b)$ is a t-conorm.

    (6 Marks)

    (b) Consider two fuzzy sets X and Y defined in the interval [0, 2], with the following respective membership functions:

$$\mu_X = \begin{cases} 1-\dfrac{x}{2} & 0 \leq x \leq 1 \\ \dfrac{1}{2} & 1 < x \leq 3 \end{cases}$$

$$\mu_Y = \frac{y}{2} \quad 0 \leq y \leq 3$$

    (i) Determine the union of X and Y using the standard fuzzy union operation given in 4(a).

    (4 Marks)

    (ii) Compute $z = x+y$ and draw the membership function μ_z.

    (10 Marks)

5. Consider an unknown system ("black box") with two inputs x_1 and x_2, and one output y as shown in Figure 1.

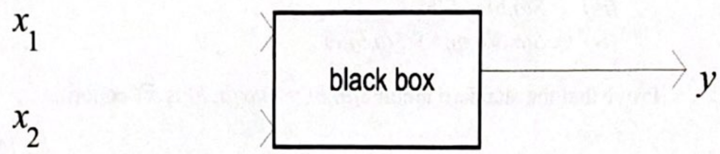


Figure 1

Suppose that inputs x_1 and x_2 are within the range X_1 = [0, 20] and X_2 = [0, 10], respectively. Two fuzzy sets S_1 and L_1 are defined in X_1 with two membership functions, μ_{S1} and μ_{L1}, respectively, as shown in Figure 2.

0,1 $\quad 1=b \quad a=-\frac{1}{15}$

15,0 $\quad 0=15a+b$

$y=-\frac{1}{15}x_1+1$

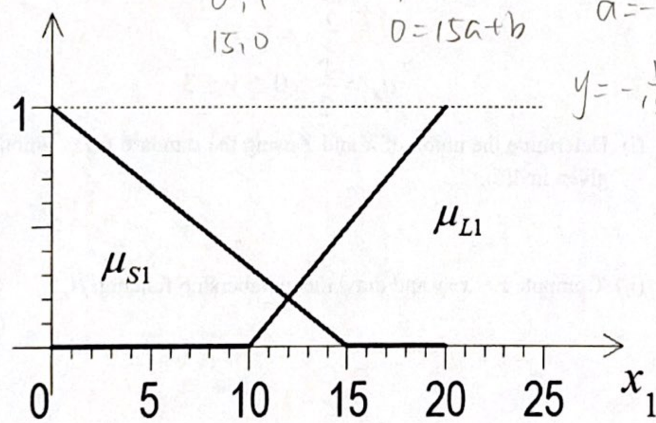


Figure 2

15, 0 $\quad 0=15a+b$

0 1 $\quad 1=b$

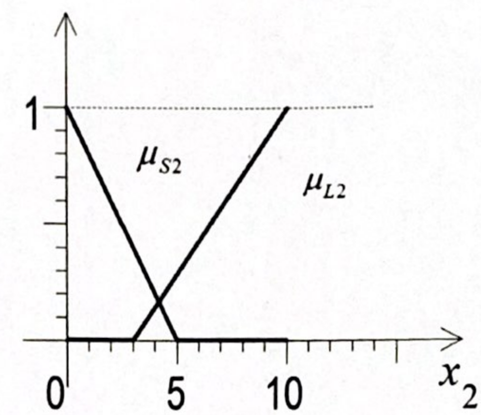
10, 0 $\quad 0=10a+b$

20, 1 $\quad 1=20a+b$

$a=\frac{1}{10} \quad b=-1$

Note: Question No. 5 continues on page 5

Similarly, X_2 has two fuzzy sets S_2 and L_2 with membership functions μ_{S2} and μ_{L2}, respectively, as shown in Figure 3.



3, 0
10, 1
0 = 3a+b
1 = 10a+b
7a = 1 a = 1/7
b = -3/7

Figure 3

(a) Determine analytical expressions of μ_{S1}, μ_{L1}, μ_{S2} and μ_{L2}. Give a suitable linguistic value to each of the 4 fuzzy sets.

(6 Marks)

(b) From experiments, the following rule base has been obtained:

Rule 1: IF x_1 is S_1 AND x_2 is L_2 THEN $y = x_1 + 4\,x_2$

Rule 2: IF x_1 is L_1 THEN $y = -6\,x_1$

(i) Based on the centre of gravity (COG) defuzzification, compute the output value y for $x_1 = 13$ and $x_2 = 4$.

(8 Marks)

(ii) Determine the output fuzzy set Y without computing its membership function.

(6 Marks)

End of Paper