Name: ________________________

Midterm Exam

Statistical Pattern Recognition

Department of Computer Engineering
Sharif University of Technology
Spring 2007

1. You must work out all the problems.
2. Show all your work in the space provided to justify your answers.
3. You may use your text book & class notes in this exam.

No	**Max**	**Score**
Problem 1		
Problem 2		
Problem 3		
Problem 4		
Problem 5 (Bonus)		
Total		

Good Luck!

1. Suppose that there are two classes, w1 and w2 and a one dimensional feature x. the likelihood functions for the two classes are given as follows:

$$p(x \mid w1) = \frac{1}{2\sqrt{2\pi}} e^{-(x+2)^2/8}; \qquad p(x \mid w2) = \frac{1}{4} e^{-|x-2|/2}$$

(a) Given P(w1)=P(w2)=0.5, calculate the minimum error rate classification decision for x= 12, x=-2, and x=4.

(b) Given P(w1)=P(w2)=0.5, find analytically the decision boundaries at decision regions.

(c) Given p(w1)=0.1 and P(w2)=0.9, find analytically the decision boundaries at decision regions.

(d) Derive a discriminant function in terms of x, P(w1) and P(w2) that is equivalent to the minimum error rate classification and does not require computing exponentials.

2.(a) Consider a random variable x having the following distribution:

$$p(x \mid \theta) = \begin{cases} 2\theta x e^{-\theta x^2} & 0 \le x, \theta > 0; \\ 0 & otherwise \end{cases}$$

Suppose that n samples are drawn independently according to $p(x \mid \theta)$. What will be the maximum likelihood estimate of θ?

3. Which of the following would be a good objective function to use instead of fisher's one and which of them not? Give 1 sentence explanations.

$$J(v) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2} + \frac{(\mu_1 - \mu_2)^2}{\sigma_2^2}$$

$$J(v) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 / \sigma_2^2}$$

$$J(v) = \frac{\sigma_1^2 * \sigma_2^2}{(\mu_1 - \mu_2)^2}$$

4. (a)In which conditions Gaussian Mixture Model (GMM) exhibit equivalent to K - means?

(b) For what sort of data does general Gaussian Mixture Model produce much better results than K-mean? Provide an example of such dataset.

(c) Suppose you had the following one-dimensional data points: 1, 4, 6, 7, 9, 13, 14, 21, 24, 28, and 29 – and you picked the points 1 and 21 as seeds for two-means clustering (i.e. k-means with k=2). What would the two clusters be after one iteration of k-means? What would the two clusters be after k-means converges? In how many steps did k-means converge?

5.(Bonus Problem) Suppose we have L classes $w_1, w_2, \ldots, w_L$ and each example feature vector x belongs to one and only one class. Suppose further that the class labeling scheme assigns a binary vector y with L components to each example with

$$y_i(x) = 1 \; if \; x \in w_i \; and \; y_j(x) = 0 \; \forall j \neq i$$

We can then write the probability of the class label vector y for a sample with features x as a multinomial distribution

$$p(y \mid x) = \prod_{i=1}^{L} \alpha_i(x)^{y_i} \qquad (1)$$

with $0 \leq \alpha_i(x) \leq 1$. For example,

$$p((0,1,0,0,0,\ldots,0) \mid x) = \alpha_2(x).$$

(a) We want to be sure that p(y|x) is properly normalized, that is

$$\sum_{\{y\}} p(y \mid x) = 1 \qquad (2)$$

where the sum is over the set of all allowed vectors y. show that this normalization condition requires that

$$\sum_{i=1}^{L} \alpha_i(x) = 1 \qquad \forall x. \qquad (3)$$

(To be clear, you should probably explicate the sum over the allowed label vectors by giving the first several terms in the sum in (2).)

(b) Suppose we have a collection of N statistically independent samples with feature vectors $x^{(a)}$ and label vectors $y^{(a)}$, a=1,2,…, L (the superscript carries the sample number). Write the likelihood of the data set

$$p(\{y^{(1)}, y^{(2)}, \ldots, y^{(N)}\} \mid \{x^{(1)}, x^{(2)}, \ldots, x^{(N)}\}, \alpha_1(x), \ldots, \alpha_L(x)) \qquad (4)$$

that follows from the likelihood for each data sample from equation (1).

(c) Show that maximizing the log-likelihood of the entire data set is equivalent to minimizing the cost function

$$\varepsilon = -\sum_{a=1}^{N} \sum_{i=1}^{L} y_i^{(a)} \ln \alpha_i(x^{(a)}). \qquad (5)$$