University of Tehran

School of Electrical and Computer Engineering

Pattern Recognition

Assignment 2

Due Date: 18th Aban

Corresponding TA:

Mehdi Bashiri – mehdi.bashiri.b@gmail.com

Aban 96

PROBLEM 1

I. Briefly describe on which condition:

   a. MAP and ML are exactly the same.
   b. Naive Bayes and MAP are equivalent.

II. Let $\{x_k\}$, k=1, 2, … , N denote independent training samples from one of the following densities. Obtain the Maximum Likelihood estimate of θ in each case.

   a. $f\,(x_k;\,\theta) = \theta\, exp(-\theta x k\,)$ $\quad x_k \geq 0\ ,\ \theta > 0$ $\quad$ *Exponential Density*
   b. $f\,(x_k;\,\theta) = \frac{x_k}{\theta^2} \exp\left(-\frac{x_k^2}{2\theta^2}\right)$ $\quad x_k \geq 0\ ,\ \theta > 0$ $\quad$ *Rayleigh Density*
   c. $f\,(x_k;\,\theta) = \sqrt{\theta}\; x_k^{\sqrt{\theta}-1}$ $\quad 0 \leq x_k \leq 1\ ,\ \theta > 0$ $\quad$ *Beta Density*

PROBLEM 2

Let x have a uniform density

$$f_x(x|\theta) \sim U(0,\theta) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \\ 0 & \text{Otherwise} \end{cases} \qquad (1)$$

a. Suppose that n samples $D = \{x_1, x_2, \ldots, x_n\}$ are drawn independently according to $f_x(x|\theta)$. Show that the maximum likelihood estimate for θ is max[D], i.e., the value of the maximum element in D.

b. Suppose that n = 5 points are drawn from the distribution and the maximum value of which happens to be $\max_k x_k = 0.6$. Plot the likelihood $f_x(D|\theta)$ in the range $0 \leq \theta \leq 1$.

Explain in words why you do not need to know the values of the other four points.

PROBLEM 3

Consider a D-dimensional Gaussian random variable x with distribution $N(x|\,\mu,\Sigma)$ in which, the covariance Σ is known and for which we wish to infer the mean μ

from a set of observations $X = \{x_1, \ldots, x_N\}$. Given a prior distribution $p(\mu) = N(\mu| \mu_0, \Sigma_0)$, find the corresponding posterior distribution $p(\mu|X)$.

PROBLEM 4

Consider four Gaussian pdf's N_1 (1.0, 0.01), N_2 (1.4, 0.13), N_3 (2.0, 0.05) and N_4 (3.3, 0.02), where two numbers indicate mean (μ_i) and variance (σ_i^2), respectively. Generate 4000 samples according to the following rule. The first sample comes from the first Gaussian pdf (N_1), the second one from N_2, the third one from N_3, and the last one from N_4. This rule repeats until all 4000 samples have been generated. The pdf underlying the random sample is modeled as the following mixture density

$$\sum_{i=1}^{I} p_i N(\mu_i \,.\, \sigma_i^2) \qquad (2)$$

a. Consider I = 4. Use the EM algorithm and the generated samples to estimate the unknown parameters, (μi, σ_i^2, *and* p_i, $i = 1, 2, \ldots, I$)
b. Repeat (a) for I = 2. Compare the result.
c. Repeat (a) for I = 3. Compare the result.
d. Repeat (a), using only 1600 generated samples. Compare the result.

PROBLEM 5

The figure below illustrates two-dimensional training samples from two classes; where, × and ● represent classes 1 and 2, respectively. Classify the test sample $(0.5, 0)^t$ using:

a. Parzen window approximation using the uniform window function

$$\varphi(u_1 . u_2) = \begin{cases} 1 & |u_i| \le 0.5 \;, i = 1,2 \\ 0 & \text{Otherwise} \end{cases} \qquad (3)$$

i. v_n is a square with a side of $h_N = 2$

ii. v_n is a circle with a radius of $r_N = 1$

b. Unbiased K Nearest Neighbor approximation with k = 3, and

i. Euclidean distance: $d(x,y) = \sqrt{\Sigma_i (x_i - y_i)^2}$

ii. City Block distance: $d(x,y) = \max_i |x_i - y_i|$

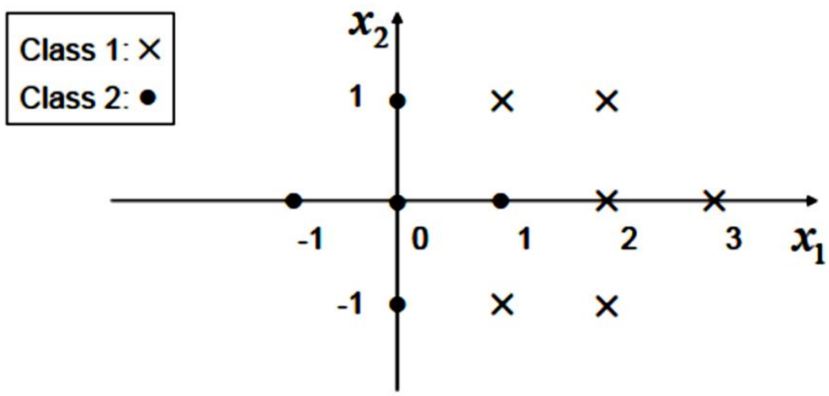

Figure 1

Hint: Any point on the border of the volume is considered an inside point.

PROBLEM 6 - HMM (25% Bonus)

Compute the probability of the observation sequences 3213 and 3212 using Forward algorithm in the following HMM (fig.2). Which one is more likely?

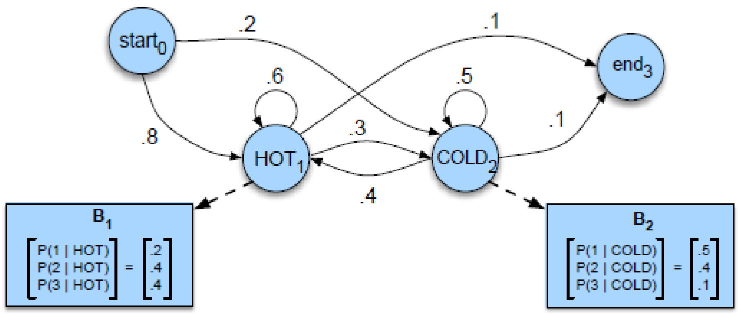

Figure 2 A hidden Markov model for relating numbers of ice creams eaten (the observations) to the weather (H or C, the hidden variables).

** Figure 2 shows a sample HMM for the ice cream task. The two hidden states (H and C) correspond to hot and cold weather, and the observations (drawn from the alphabet O = {1,2,3}) correspond to the number of ice creams eaten by Jason on a given day.

PROBLEM 7 – HMM (25% Bonus)

I. Show that for a Markov Chain: $P(q_t|q_{t+1}, \ldots, q_T) = \mathrm{P}(q_t|\, q_{t+1})$.

II. Consider the following 2-state Hidden Markov Models where both states have two possible output symbols A and B.

Model 1:
Transition probabilities: a_{11}=0.7, a_{12}=0.3, a_{21}=0.0, a_{22}=1.0
(a_{ij} is the probability of going from state i to state j)
Output probabilities: $b_1(A) = 0.8$, $b_1(B) = 0.2$, $b_2(A) = 0.4$, $b_2(B) = 0.6$.
Initial probabilities: π_1=0.5, π_2=0.5 .

Model 2:
Transition probabilities: a_{11}=0.6, a_{12}=0.4, a_{21}=0.0, a_{22}=1.0 .
Output probabilities: $b_1(A) = 0.9$, $b_1(B) = 0.1$, $b_2(A) = 0.3$, $b_2(B) = 0.7$.
Initial probabilities: π_1=0.4, π_2=0.6 .

a. Sketch the state diagram for two models.

b. Which model is more likely to produce the observation sequence {A, B, B}?

Hint: Information you need for solving the last two problems can be found here:

https://web.stanford.edu/~jurafsky/slp3/9.pdf