MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Media Arts and Sciences Section

Course 4.902: Pattern Recognition and Analysis   Assigned: 3/11/93

Problem Set No. 6   Due: 4/1/93

QUIZ 1 REMINDER

> Quiz 1 will be held during class time on Tuesday, March 16.
>
> Material from the following will be covered: Class lectures through March 4; Therrien Chapters 1-6; Problem Sets 1-5.

READING ASSIGNMENT: Therrien, Chapter 8 and Sections 9.1 and 9.3

This problem set consists of five problems. Remember that you are *not* allowed to look at old problem sets, but collaboration is allowed and encouraged.

Feel free to use Matlab or any software package of your choice on any of these problems.

Problem 1: (10 points)

It is suspected that a certain six-sided die is loaded (unfair). By rolling the die many times, we wish to estimate the probability of each of the six outcomes. Let $\{P_i\},\ i = 1, \ldots, 6$ denote the unknown probabilities.

Suppose the die is rolled N times. If the probabilities to be estimated were known, then the probability of observing exactly y_1 1's, y_2 2's, ..., and y_6 6's in the N rolls would be:

$$P(y_1, \ldots, y_6\ ;\ P_1, \ldots, P_6) = \frac{N!}{y_1! y_2! \cdots y_6!} P_1^{y_1} P_2^{y_2} \cdots P_6^{y_6}.$$

This distribution is called the *multinomial distribution.*

a) Suppose exactly y_i occurrences of i are observed in the N rolls, where $i = 1, \ldots, 6$. Find the maximum likelihood estimates $\{\hat{P}_i\}$ for the probabilities $\{P_i\},\ i = 1, \ldots, 6$ in terms of the y_i's.

   Hint: Take the logarithm of the likelihood function, then use Lagrange multipliers.

b) Using your analysis from part (b) and the fact that the maximum likelihood estimate is consistent and asymptotically unbiased, show that for any two probability mass functions $\{P_i\}$ and $\{\hat{P}_i\}$,

$$\sum_i P_i \log \frac{P_i}{\hat{P}_i} \geq 0.$$

The quantity on the left-hand side is the *Information Divergence* or *Kullback Leibler distance* between P and $\hat{P}$ for the discrete case, and the inequality corresponds to the one you derived (in a different way) in Problem Set 3, Problem 1 (a) for continuous densities.

Problem 2: (10 points)

Therrien, Problem 7.3.

Problem 3: (10 points)

Therrien, Problem 7.5.

Problem 4: (10 points)

Construct the 1-D Parzen estimate for a density p_y given the samples

$$y^{(i)}: \; 0.2,\ 2.2,\ 2.7,\ 3.3,\ 3.5,\ 4.3,\ 4.4,\ 4.5,\ 4.8,\ 5.2,\ 5.5$$

for each of the following kernels (as shown in Therrien, Figure 8.3). Find reasonable choices of parameters for each kernel.

a) Rectangular.
b) Gaussian.

Problem 5: (60 points)

Choose one of the following real problem scenarios and consider how pattern recognition may be used to solve it.

Scenario 1: A cattle rancher wants to have a non-invasive system of recognizing cows by the patterns on their backs. He has a few hundred cows, and they will walk under an archway where a photo is taken of their back from a camera mounted above. You are given only the image data, and you need to tell him the name of this cow. (The rancher wants to keep track of individual data on each cow, their health shots, etc.)

Scenario 2: The post office wants to speed up the delivery of mail by automating the verification of postage stamps on each letter. A camera will snap a photo of each letter as it moves by on a conveyor belt. Given this image, you must find the stamp(s), read their values, and output the amount of postage on the letter. (Assume your output will be put into another system for comparison with the weight of the letter.)

You should consider (but not feel limited to) ideas from matched filtering, eigenvectors, decision theory, and neural nets. Please limit your response to 3 pages. Make sure your solution includes the following:

(a) Describe possible features you would collect, and any training processes you might use.
(b) How will you compare features and make decisions?
(c) What assumptions are you making?
(d) How would you test your system?
(e) Name some conditions where you expect your system would fail.
(f) Suggest some ways these failures might be handled.

Please be as clear as you can, showing mathematically and with diagrams how your solution might work.