SD 372 Intro. to Pattern Recognition Feb. 13, 2001

Midterm Examination

Professor Paul Fieguth

Time:	75 Minutes
Grading:	The exam is out of 40. The grade breakdown is shown in the margins.
Aids Permitted:	*One* 8.5x11" page, *No* calculator.
Advice:	Read problems carefully before jumping in to calculations. Well-drawn sketches / diagrams can be very helpful. If you can't answer a question, you *will* receive part marks for relevant statements, insights, or sketches. Tell me what you know! Avoid long / verbose answers; be succinct.

Bonus Question

Do **NOT** waste your time here unless you're finished with the exam ...

Suppose we take a regular coin and flip it ∞ times.

What is the probability that 10^6 heads appear in a row?

What is the probability that, in the whole sequence, 10^6 heads appear in a row an infinite number of times?

Question 1.

Give answers to each of the following; note that each part of the question is independent of the other parts.

a) Draw a sketch showing the unit standard-deviation shape for the following two clusters:

$$\underline{\mu}_1 = \begin{bmatrix} 0 \\ -2 \end{bmatrix} \quad \Sigma_1 = \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix} \qquad \underline{\mu}_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad \Sigma_2 = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

b) Suppose we have a cluster in *three* dimensions:

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \underline{\mu}_x = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \Sigma_x = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 5 & -3 \\ 1 & -3 & 4 \end{bmatrix}$$

Clearly we can't sketch the three-dimensional shape, however we *do* know how to sketch in two dimensions. Sketch the shape of the cluster in the two-dimensional feature space (x_1, x_3).

c) In the classroom and in the lab we have generally stuck to one- and two-dimensional problems. *Briefly,* what are some of the challenges associated with applying Pattern Recognition to high-dimensional problems?

d) State Bayes' rule (there are many ways of writing it, just give me one).

Why is Bayes' rule so important in Pattern Rec? That is, what role does it play? *Be brief!*

e) A substantial, current socio-environmental problem is urban sprawl. One of the biggest problems is even knowing *how much* sprawl has taken place.

There has recently been some interest in quantifying sprawl based on satellite measurements. What sorts of features might a satellite be able to measure that would allow you to set up a Pattern Recognition problem to discriminate between the classes C_{city} and $C_{country}$?

Keep your answer brief.

Question 2.

Human neurons communicate by firing electric signals repeatedly over time. For certain neurons the time x between successive firings is exponential:

$$p(x|C_1) = \begin{cases} e^{-t} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

In other nerve cells, there is a certain recovery time which is required before the cell can fire again, so the time interval is delayed:

$$p(x|C_2) = \begin{cases} 2e^{-2(t-1)} & x \geq 1 \\ 0 & x < 1 \end{cases}$$

We can try to classify a neuron into one of these two classes by measuring the inter-firing time x. *Unless otherwise stated,* assume that the two types of neurons are equally likely:
$P(C_1) = P(C_2) = 0.5$.

a) Draw a sketch of the two classes.

b) Find the mean and variance for each of the two classes.

c) Write down the MAP classifier *as a function of* $P(C_1)$.

d) Find the MED and MICD classifiers.

e) Assume that the two classes are equally likely. Find expressions for the probabilities of error $P_{MED}(\epsilon)$ and $P_{MAP}(\epsilon)$. Which error is smaller?

f) Carefully redraw your sketch from (a), but now on your sketch compare $P_{MED}(\epsilon)$, $P_{MICD}(\epsilon)$, and $P_{MAP}(\epsilon)$. If you prefer, you may find it neater to draw three separate sketches – your choice.

   Based on your sketch, approximately how similar or different are the performances of the three classifiers?