GI12/4C59 – Homework #3 (Due 12am, November 29, 2005)

Aim: To get familiarity with the basic concepts of eigenvalue problems, probabilistic inference, and noisy channels. Presentation, clarity, and synthesis of exposition will be taken into account in the assessment of these exercises.

This document is available from
http://www.cs.ucl.ac.uk/staff/Rob.Smith/Internal/information_theory.htm

1. **[30 pts]** The standard eigenvalue problem is intimately involved in many aspects of data compression, learning, dynamical systems theory, etc. Therefore, a brief consideration of such problems will be of use.

   Consider a second-order, homogeneous dynamic system of the form:

   $$\frac{d^2 y}{dt^2} + 2\xi\omega_n \frac{dy}{dt} + \omega_n^2 y = 0$$

   (a) Place the system in first order form.

   (b) Solve for y as a function of time (you can do this in either first-order or the original, second-order form above).

   (c) Based on your solution, comment on the importance of the value ranges of the parameter ξ (often called the *damping ratio*).

2. **[40 pts]**: There are three cards. One card is white on both sides, one is black on both sides, and one is black on one side and white on the other. The three cards are shuffled and their orientations randomized. One card is drawn and placed on the table. The upper face is black.

   (a) What is the color of its lower face?

   (b) Does seeing the top face of the card on the table convey *information* about the color of the bottom face? Discuss the *information contents* and *entropies* in this situation. Let the value of the upper face's color be *u* and the value of the lower face's color be *l*. Imagine that we draw a random card and learn both *u* and *l*. What is the entropy of *u*, *H(U)*? What is the entropy of *l*, *H(L)*? What is the mutual information between *U* and *L*, *I(U;L)*?

3. **[30 pts]** Given *the state transition probability matrix* (that is, the probability of output *y* given input *x* for a channel with 5 input characters and 10 output characters is:

$$\begin{bmatrix} 0.25 & 0 & 0 & 0 & 0.25 \\ 0.25 & 0 & 0 & 0 & 0.25 \\ 0.25 & 0.25 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0 & 0 & 0 \\ 0 & 0.25 & 0.25 & 0 & 0 \\ 0 & 0.25 & 0.25 & 0 & 0 \\ 0 & 0 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0.25 & 0 \\ 0 & 0 & 0 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0.25 & 0.25 \end{bmatrix}$$

What is the channel capacity?

Model Solutions:

1.

(a)

$$\frac{d^2y}{dt^2}+2\xi\omega_n\frac{dy}{dt}+\omega_n^2y=0$$
$$\frac{dy}{dt}=v$$
$$\frac{dv}{dt}=-2\xi\omega_n v-\omega_n^2 y$$
$$\frac{d}{dt}\begin{bmatrix} y \\ v \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\xi\omega_n \end{bmatrix}\begin{bmatrix} y \\ v \end{bmatrix}$$

(b)

$$\frac{d^2y}{dt^2}+2\xi\omega_n\frac{dy}{dt}+\omega_n^2y=0$$
$$y=C\exp(\lambda t)$$
$$\frac{dy}{dt}=C\lambda\exp(\lambda t)$$
$$\frac{d^2y}{dt^2}=C\lambda^2\exp(\lambda t)$$
$$C\lambda^2\exp(\lambda t)+2\xi\omega_n C\lambda\exp(\lambda t)+\omega_n^2 C\exp(\lambda t)=0$$

Neglecting the trivial solution

$$\lambda^2+2\xi\omega_n\lambda+\omega_n^2=0$$
$$\lambda=\frac{-2\xi\omega_n\pm\sqrt{(2\xi\omega_n)^2-4\omega_n^2}}{2}$$
$$\lambda=-\omega_n\left[\xi\pm\sqrt{\xi^2-1}\right]$$
$$y=C_1\exp\left(-\omega_n\left[\xi+\sqrt{\xi^2-1}\right]t\right)+C_2\exp\left(-\omega_n\left[\xi-\sqrt{\xi^2-1}\right]t\right)$$

(c)

The parameter ξ clearly determines stability, since for negative values system response *y* will grow without bound.

In addition (and perhaps more importantly), consider the following fact, derived from Euler's Identity:

$\exp((r\pm\omega i)t)=\exp(rt)[2\cos\omega t]$,

and note that in the solution, the parameter ξ completely determines whether there is an imaginary part to the eigenvalues.

Via these facts, if there is an imaginary part ($-1>\xi>1$) then the system oscillates towards its equilibrium (or towards infinity, in the unstable $\xi<0$ case). Otherwise, the system does not oscillate, but approaches its equilibrium without overshooting (or infinity, in the unstable $\xi<0$ case). There are persistent oscillations for $\xi=0$ (but I didn't require you to say this last item for full credit).

2.

(a) There is, of course, no way to know the color of the lower face with certainty in this situation. However, we can calculate the probabilities that the card on the table has black on both sides (BB) or white on the lower face (BW):

$$P(BB \mid U=B) = \frac{P(U=B \mid BB)P(BB)}{P(U=B)} = \frac{(1)(1/3)}{1/2} = 2/3$$

$$P(BW \mid U=B) = \frac{P(U=B \mid BW)P(BW)}{P(U=B)} = \frac{(1/2)(1/3)}{1/2} = 1/3$$

We can note at this point that seeing the color of the upper face changed the probability of the lower face's color (which beforehand was evenly divided between the two colors). Thus, are instinct for part (b) is that information has been conveyed.

(b)

$$H(U) = H(L) = -2(1/2)\log_2(1/2) = 1$$

$$H(U,L) = -\sum_U \sum_L P(U,L)\log_2 P(U,L)$$

$$= -2(1/3)\log_2(1/3) - 2(1/6)\log_2(1/6)$$

$$= 1.9183 \text{ bits}$$

$$I(U;L) = H(U) + H(L) - H(U,L) = 1 + 1 - 1.9183 = 0.0817 \text{ bits}$$

3. We find the channel capacity by maximizing the mutual information between input and output, with respect to the probability distribution over inputs:

$$C = \max_{P(X)} I(X;Y) = \max_{P(X)}\left(H(Y) - H(Y \mid X)\right)$$

$$H(Y \mid X) = \sum_{x \in X} P(x) H(Y \mid X = x)$$

$$H(Y \mid X = x) = -\sum_{y \in Y} P(Y \mid X = x)\log_2 P(Y \mid X = x) = -4\frac{1}{4}\log_2\frac{1}{4} = 2 \quad (\forall x)$$

$$H(Y \mid X) = 2\sum_{x \in X} P(x) = 2$$

$$C = \max_{P(X)}\left(H(Y) - 2\right)$$

So, for this particular problem, we get maximum channel capacity when we maximize the entropy of the output, *by varying the probability distribution of the input.* We know that we will get maximum entropy for a uniform distribution. Since the columns of the transition matrix are all permutations of one another, all symbols Y have identical distributions with respect to X. Therefore, a uniform distribution of the input will yield a uniform distribution of the output, and:

$$\max_{P(X)}\left(H(Y)\right) = -\sum_{y \in Y} P(Y)\log_2 P(Y) = -10(1/10)\log_2(1/10) = 3.32$$

$$C = 3.32 - 2 = 1.32 \text{ bits}$$