Problem Set 1

Review of Linear Algebra and Probability

Prof. Jim Rehg
CS 7636 Computational Perception
College of Computing
Georgia Institute of Technology

January 17, 2003

Solutions must be turned in by 5 pm January 20, 2003. Matlab will be helpful for some parts. Hand-written answers are fine as long as they are legible and organized. Please show all work. I've tried to leave space on this handout for your solutions.

Vectors and matrices are typeset in bold. A vector $\mathbf{v}$ is a column vector; $\mathbf{v}^T$ is a row vector where "T" denotes the transpose. $P(\cdot)$ denotes a probability density function (PDF), which sums or integrates to one.

1. [40 points]

An eager young professor who is heading back into the lab late at night is trying to guess whether his graduate student is working on her paper or not. He has noticed that whenever this particular student is working, her car is in the parking lot and loud music is playing in the lab. Let W, C, M be three binary random variables corresponding to the events "student Working on paper", "Car in parking lot", and "loud Music playing". Through careful observation, the professor has determined the following conditional probabilities:

$$P(C|W,M) \; = \; P(C|W) \; = \; \begin{array}{c|cc} & W=1 & W=0 \\ \hline C=1 & 0.8 & 0.1 \\ C=0 & 0.2 & 0.9 \end{array} \; = \; P(M|W)$$

In other words, C and M are conditionally independent given W, and have the same distribution e.g. $P(C=1|W=1) = P(M=1|W=1) = 0.8$. The professor guesses that $P(W=1) = 0.8$ (she is a good student).

(a) [15 points]
Compute the initial distributions $P(C)$ and $P(M)$, which reflect the state of knowledge before the professor arrives on campus. Hint: First compute the joint probabilities $P(C, W)$ and $P(M, W)$ and then marginalize (sum) over W.

(b) [10 points]
Upon driving into the parking lot, the professor notices his student's car. Update the distribution over W to reflect this new information (i.e. calculate $P(W|C = 1)$). Hint: Use Bayes Rule.

(c) [10 points]
Calculate the probability that loud music is playing in the lab. Hint: Calculate the updated $P(M, W)$.

(d) [5 points]
Upon reaching his office, the professor remembers that the conference deadline was yesterday! What is $P(M)$?

2. [40 points]

(a) [10 points]
Write the probability density function for a gaussian random vector $\mathbf{x}$ with mean μ and covariance $\mathbf{\Sigma}$ (i.e. the vector form of the normal density).

(b) [20 points]
Someone has given you a large database of random vectors of length 2, distributed according to a gaussian density with zero mean and unit covariance ($\mathbf{\Sigma} = \mathbf{I}$). For a simulation, you need random vectors distributed with the following mean and covariance:

$$\mu = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \qquad \mathbf{\Sigma} = \begin{bmatrix} 3/8 & -1/8 \\ -1/8 & 3/8 \end{bmatrix}$$

Give an algorithm (with sufficient detail to implement it) for constructing the desired random vectors from the vectors in your database.

(c) [10 points]
Sketch the equiprobability contours of the desired gaussian density in part (b).

3. [20 points]

(a) [10 points]
x is a scalar gaussian random variable with mean μ and variance σ_x^2. Measurements of x are corrupted by an independent additive gaussian noise source n. The measurement model is $y = x + n$. n has zero mean and variance σ_n^2. Write the following (Note: These are not trivial to derive, and it's fine if you look them up):

- An expression for $p(x|y)$, the conditional density which encapsulates our knowledge of x as a function of observations y.

- The estimate $\hat{X}$ for a given $y = Y$ that minimizes $E\{\|x - \hat{X}\|^2 | y = Y\}$, where $E\{\cdot | y = Y\}$ denotes the conditional expectation with respect to x. This estimate goes under various names, including least squares, minimum mean square, and minimum variance.

(b) [10 points]
Let s be an unknown random variable. A complex random process involving s generates a zero mean measurement vector $\mathbf{y}$ of length 3. Through careful experimentation you determine that:

$$E\{\mathbf{y}\mathbf{y}^T\} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 5 \end{bmatrix} \quad \text{and} \quad E\{s\mathbf{y}\} = \begin{bmatrix} 2 \\ 5 \\ 12 \end{bmatrix}$$

A program left behind by one of your previous graduate students makes it easy to compute estimates of the form $\hat{s} = \mathbf{a}^T\mathbf{y}$ where $\mathbf{a}$ is an arbitrary constant vector. Calculate the value of $\mathbf{a}$ which will yield the minimum mean square estimate of s (i.e. minimizing $E\{\|s - \hat{s}\|^2\}$ where the expectation is over all possible $s, \mathbf{y}$). Identify the equations you have solved.