ECE 438 Homework 10, due in class Friday, 11/5/2004

Problem 1.MAXIMUM LIKELIHOOD (ML) ESTIMATION, TAKE 1.
Recall that, for the problem of estimating a deterministic parameter vector $\mathbf{x}$ from an observation of a random vector $\mathbf{Y}$, the *likelihood function* $L(\mathbf{y}, \mathbf{x})$ is defined to be the conditional PDF of the observed random vector, given the value of the parameter vector $\mathbf{x}$:

$$L(\mathbf{y}, \mathbf{x}) \stackrel{def}{=} f_{\mathbf{Y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}).$$

Suppose that

$$\mathbf{Y} = \mathbf{x} + \mathbf{W},$$

where $\mathbf{Y}$, $\mathbf{x}$, and $\mathbf{W}$ are 2-dimensional vectors.

- $\mathbf{x} = (x(1), x(2))^T$ is an unknown, non-random vector, such that

$$|x(1)| + |x(2)| \leq 1 \tag{1}$$

- $\mathbf{W}$ is a vector of two zero-mean, statistically independent, identically distributed, Gaussian random variables with variance σ^2.

(a) Find an expression for the likelihood function.

(b) Given an observation $\mathbf{y}$ of $\mathbf{Y}$, find the estimate $\hat{\mathbf{x}}(\mathbf{y})$ which maximizes the likelihood function $L(\mathbf{y}, \mathbf{x})$:

$$L(\mathbf{y}, \hat{\mathbf{x}}(\mathbf{y})) = \max_{\text{over all } \mathbf{x} \text{ satisfying Eq. (1)}} L(\mathbf{y}, \mathbf{x}).$$

(**Hint.** Maximizing $\log L(\mathbf{y}, \mathbf{x})$ is equivalent to maximizing $L(\mathbf{y}, \mathbf{x})$, but is more tractable.) Provide a geometric interpretation for your result–i.e. interpret $\hat{\mathbf{x}}(\mathbf{y})$ as the closest point in a certain set to a certain point.

(c) Find $\hat{\mathbf{x}}(\mathbf{y})$ if

(i) $\mathbf{y} = (0, 0)^T$;

(ii) $\mathbf{y} = (3, 3)^T$;

(iii) $\mathbf{y} = (1000, 1)^T$.

Problem 2. ML ESTIMATION, TAKE 2.
Suppose that

$$\mathbf{Y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \mathbf{W},$$

where $\mathbf{Y}$, $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{W}$ are N-dimensional vectors.

- $\mathbf{x}_1$ and $\mathbf{x}_2$ are known, non-random vectors, with $\|\mathbf{x}_1\| = \|\mathbf{x}_2\| = 1$, and $\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = 0$.
- $\mathbf{W}$ is a vector of zero-mean, statistically independent, identically distributed, Gaussian random variables.

- a_1 and a_2 are unknown real numbers, to be estimated from an observation of $\mathbf{Y}$.

Given an observed value $\mathbf{y}$ of $\mathbf{Y}$, find estimates $\hat{a}_1$, $\hat{a}_2$ of a_1, a_2, respectively, which maximize the likelihood function:

$$L(\mathbf{y}, a_1, a_2) = f_{\mathbf{Y}|a_1,a_2}(\mathbf{y}|a_1, a_2).$$

(**Hint.** Write out the logarithm of the likelihood function and simplify. You should see that maximizing the likelihood is equivalent to minimizing the distance from $\mathbf{y}$ to $a_1\mathbf{x}_1 + a_2\mathbf{x}_2$.)

Problem 3. Suppose that we observe N iid random variables $X(1), \ldots, X(N)$. The PDF of each of these random variables is:

$$f_X(x) = \begin{cases} 1, & \text{if } r < x \leq r+1, \\ 0, & \text{otherwise.} \end{cases}$$

We do not know the value of r.

(a) Note that $E[X(n)] = r + 0.5$, and therefore we may estimate r as follows:

$$\hat{r}_N = \frac{1}{N}\sum_{n=1}^{N} X(n) - \frac{1}{2}.$$

(i) Find $E[\hat{r}_N]$. Is $\hat{r}_N$ an unbiased estimate of r?

(ii) Find $Var(X(n))$.

(iii) Find $Var(\hat{r}_N)$.

(iv) For $N = 48$, what is (approximately) the probability that our estimate $\hat{r}_{48}$ is within ± 0.01 of the true value of r? Within ± 0.05 of the true value? Use the central limit theorem and a table of values of a Gaussian CDF.

(b) We can also use the largest of our N experimental values as our estimate of $r + 1$:

$$\hat{r}'_N = \max(X(1), X(2), \ldots, X(N)) - 1.$$

(i) Find the CDF $F_X(x)$.

(ii) Find the CDF $F_{\hat{r}'_N}(x)$ of $\hat{r}'_N$, by using the fact that

$$\mathbf{P}(\max(X(1), X(2), \ldots, X(N)) - 1 \leq x) =$$
$$\mathbf{P}(X(1) \leq x + 1 \text{ and } X(2) \leq x + 1 \text{ and } \ldots \text{ and } X(N) \leq x + 1).$$

(iii) For $N = 48$, what is the probability that our estimate $\hat{r}'_{48}$ is within $(+0, -0.02)$ of the true value of r? Within $(+0, -0.10)$ of the true value? Compare with your results from (a)(iv).

(iv) Find the PDF of $\hat{r}'_N$.

(v) Find $E[\hat{r}'_N]$. Is $\hat{r}'_N$ unbiased? Is it asymptotically unbiased? (I.e., is $\lim_{N\to\infty} E[\hat{r}'_N] = r$?)

(vi) Find $Var(\hat{r}'_N)$. As $N \to \infty$, does it approach zero faster or slower than $Var(\hat{r}_N)$?

(c) In Matlab, use command `rand` to generate $X(1), \ldots, X(400)$ for $r = 2$. Write a function or a script to generate $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_{400}$, and plot $\hat{r}_n$ as a function of n. Write another function or script to generate $\hat{r}'_1, \hat{r}'_2, \ldots, \hat{r}'_{400}$, and plot $\hat{r}'_n$ as a function of n. Where do the plots converge and why? Argue why one of these two estimators may be preferable. (It may be convenient to plot the two sequences on the same plot, by using command `hold on`.)