Detection and Modulation Theory: Additional Homework

Erik G. Larsson

May 18, 2008

Kay-II-3.14. Design a suitable hypothesis test for the problem

$$\begin{cases} H_0 : \boldsymbol{x} \sim N\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \\ H_1 : \boldsymbol{x} \sim N\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right) \end{cases}$$

where ρ is given. Comment on the special case when $\rho \to 0$.

EL1. (After discussion by MacKay.) You are given a coin, and the task is to determine whether the coin is fair (H_0) or not (H_1).

a) In the first experiment you decide to flip the coin a fixed number (n) of times and record how many heads (h) and how many tails (t) you got ($n = h + t$). You then compute the probability p_e that *if the coin were fair, you would obtain an outcome as extreme, or more extreme, than this*. If p_e is less than a predefined, fixed "plausibility" threshold p_0 (take 5% in this problem) than you determine the coin is biased (H_1), otherwise you say it is fair (H_0).

Suppose you decide to take $n = 12$ (flip the coin 12 times) and you obtain the sequence `hhhthhhhthht`. Do you decide on H_0 or H_1?

b) In a different experiment you decide instead to flip the coin until a fixed number, say t, of tails occur. You then record how many times (n) you had to flip the coin for this to happen. As before you then compute the probability p_e that if the coin were fair, you would obtain an outcome that is at least as extreme as this, and decide on H_1 if p_e is less than the level 5%.

Suppose you decide to take $t = 3$ and obtained the string above, `hhhthhhhthht`, i.e. the 3rd tail occurred after $n = 12$ tosses. Do you decide on H_0 or H_1?

c) Do the experiments (a) and (b) lead to the same conclusion? If not, why? Which experiment (a) or (b) do you think is most "fair"? Which one do you prefer and why?

d) Design a Bayesian test and compute the a posteriori likelihood ratio explicitly. You may make any assumptions you wish, e.g., as in the lecture.

Hints: In a) you need to consider the distribution of t, for n fixed. In b), you need to consider the distribution of n, for t fixed.

EL2. Consider the BI-AGN channel with the following signal constellation and bit-symbol mapping

$b_1 = 0$, $b_2 = 0$: $s = -3$
$b_1 = 0$, $b_2 = 1$: $s = -1$
$b_1 = 1$,$b_2 = 1$: $s = 1$
$b_1 = 1$, $b_2 = 0$: $s = 3$

(This mapping is called Gray mapping in communications.)

Suppose $r = s + e$ where $e \sim N(0, \sigma)$, σ known and b_1, b_2 independent. Suppose $\sigma^2 = 1$.

a) Write up an expression for LLR$(b_1|r)$ and LLR$(b_2|r)$.

b) Suppose

$$P(b_1 = 0) = P(b_1 = 1)$$

and

$$P(b_2 = 0) = P(b_2 = 1)$$

Draw LLR$(b_1|r)$ and LLR$(b_2|r)$ as functions of r. Carefully discuss the result.

c) Repeat the experiment for

$$P(b_1 = 0) = 1000 \cdot P(b_1 = 1)$$

and

$$P(b_2 = 0) = P(b_2 = 1)$$

How does the result for $\text{LLR}(b_1|r)$ change? How does the result for $\text{LLR}(b_2|r)$ change?

We assumed b_1, b_2 are independent before r is observed. Are they still independent after r was observed?

Kay-I-3.9 Consider

$$r_1 = a + e_1$$

$$r_2 = a + e_2$$

and suppose e_1, e_2 are jointly Gaussian with known variance σ^2 and known correlation coefficient ρ.

Compute the CRB for the estimation of a based on r_1, r_2 and sketch it as a function of ρ. Discuss the result.

EL3. Consider ML estimation (deterministic parameter vector) with "nuisance parameters", that is, we are only interested in estimating a subset of the unknown parameters. More precisely, if the likelihood function is $p(\boldsymbol{r}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ where $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are unknown (vector parameters) we are only interested in estimating $\boldsymbol{\theta}_1$, whereas $\boldsymbol{\theta}_2$ is an undesired nuisance parameter in which we are not particularly interested.

a) One way to deal with nuisance parameters is to form the "concentrated" ML estimate

$$\hat{\boldsymbol{\theta}}_1 = \text{argmax}_{\boldsymbol{\theta}_1} \bar{p}(\boldsymbol{r}|\boldsymbol{\theta}_1)$$

where

$$\bar{p}(\boldsymbol{r}|\boldsymbol{\theta}_1) \triangleq \max_{\boldsymbol{\theta}_2} p(\boldsymbol{r}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$$

Consider the partially linear model

$$\boldsymbol{r} = \boldsymbol{A}(\boldsymbol{\theta}_1)\boldsymbol{\theta}_2 + \boldsymbol{e}$$

where $\boldsymbol{e} \sim N(\boldsymbol{0}, \boldsymbol{Q})$ for some positive definite $\boldsymbol{Q}$. Here $A(\boldsymbol{\theta}_1)$ indicates that $\boldsymbol{A}$ is a nonlinear matrix function of $\boldsymbol{\theta}_1$. Derive the concentrated ML estimate of $\boldsymbol{\theta}_1$. You can assume that $\boldsymbol{A}(\boldsymbol{\theta}_1)$ has full column rank for all $\boldsymbol{\theta}_1$.

b) As alternatives to "concentrating" the likelihood function (forming $\bar{p}(\cdot)$), an alternative approach is to consider $\boldsymbol{\theta}_2$ random and eliminate it by forming

$$\hat{\boldsymbol{\theta}}_1 = \arg\max_{\boldsymbol{\theta}_1} \bar{p}(\boldsymbol{r}|\boldsymbol{\theta}_1)$$

where

$$\bar{p}(\boldsymbol{r}|\boldsymbol{\theta}_1) \triangleq E_{\boldsymbol{\theta}_2}[p(\boldsymbol{r}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)] = \int p(\boldsymbol{r}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)p(\boldsymbol{\theta}_2)d\boldsymbol{\theta}_2$$

Consider again the model $\boldsymbol{r} = \boldsymbol{A}(\boldsymbol{\theta}_1)\boldsymbol{\theta}_2 + \boldsymbol{e}$ and suppose it is known a priori that $\boldsymbol{\theta}_2 \sim N(\boldsymbol{\mu}, \boldsymbol{\Psi})$ for some positive definite $\boldsymbol{\Psi}$. What is the resulting estimate of $\boldsymbol{\theta}$?

c) Comment on how the differences/similarities between the result in (a) and (b).

EL4. Consider the scalar Gaussian model where r_i i.i.d.

$$r_i \sim N(\mu(\theta), \sigma(\theta))$$

for $i = 1, ..., N$. Derive the scalar version of Slepian-Bang's formula:

$$I(\theta) = \frac{N}{\sigma^2(\theta)}\left(\frac{d\mu(\theta)}{d\theta}\right)^2 + \frac{N}{2}\frac{1}{\sigma^4(\theta)}\left(\frac{d\sigma^2(\theta)}{d\theta}\right)^2$$

EL5. Prove the formulas for GLRT under linear Gaussian model presented in class (left column only, i.e. the case $\boldsymbol{\theta} = \mathbf{0}$ vs $\boldsymbol{\theta} \neq \mathbf{0}$).

EL6. Consider

$$H_0 : \begin{cases} x_1^i = \alpha\cos(\phi) + e_1^i \\ x_1^q = \alpha\sin(\phi) + e_1^q \\ x_2^i = \alpha\cos(\phi) + e_2^i \\ x_2^q = \alpha\sin(\phi) + e_2^q \end{cases}$$

$$H_1 : \begin{cases} x_1^i = \alpha\cos(\phi) + e_1^i \\ x_1^q = \alpha\sin(\phi) + e_1^q \\ x_2^i = -\alpha\cos(\phi) + e_2^i \\ x_2^q = -\alpha\sin(\phi) + e_2^q \end{cases}$$

You want to discriminate between H_0 and H_1 based on $x_1^i, x_1^q, x_2^i, x_2^q$. The noises $e_\times^\times$ are i.i.d. $N(0, \sigma)$ where σ is known. The gain α and phase ϕ are unknown.

a) Explain why the problem is a simple model for demodulation of differential BPSK transmitted over a narrowband channel with amplitude α and phase ϕ, where ϕ incorporates the phase of the previous symbol.

b) Suppose ϕ is unknown, uniformly distributed over $[-\pi, \pi]$ but α is known. Construct the optimal Bayesian test.

c) Under the assumptions made in (b), compute

$$\log\left(\frac{P(H_1|x_\times^\times)}{P(H_0|x_\times^\times)}\right)$$

What assumptions do you need to compute this quantity? Why is this quantity important?

d) (optional) Comment on the case when neither ϕ nor α is known and suggest a decision rule.

Hint 1. In the calculations it can be helpful to work with complex numbers. However, there is no need to work with complex random numbers.

Hint 2. You may not be able evaluate all integrals in closed form. (Use Bessel functions if you want.)

EL7. (Expansion in discrete time.)

Consider a seris of samples $\{x_k = x(kT)\}_{k=-N}^{N}$ of a zero mean stationary continuous-time random process $x(t)$. Let $\boldsymbol{x} = [x_{-N}, ..., x_N]^T$, $r[k] = E[x_n x_{n-k}]$ and let $\boldsymbol{R} = E[\boldsymbol{x}\boldsymbol{x}^T]$ (i.e., $R_{ij} = r[i-j]$).

Let

$$Y(\omega) = \frac{1}{\sqrt{2N+1}} \sum_{k=-N}^{N} x_k e^{-j\omega n}$$

be the Fourier transform of x_k.

a) Compute an expression for $E[Y(2\pi n/(2N+1))Y^*(2\pi k/(2N+1))]$. Consider the limit when $N \to \infty$. Show that asymptotically $Y(2\pi n/(2N+1)), Y(2\pi k/(2N+1))$ are uncorrelated for $k \neq n$ and compute their variance. Explain how their variance relates to the Fourier transform of $r[n]$.

For this you can assume covariance function $r[k]$ decays exponentially fast, i.e. there is a positive constant α such that $|r[k]| = O(e^{-\alpha|k|})$

b) Argue (heuristically) that for large N, the eigenvectors to $\boldsymbol{R}$ are Vandermonde vectors of the form $[e^{-jkN2\pi/(2N+1)}, ..., 1, ..., e^{jkN2\pi/(2N+1)}]^T$ and the eigenvalues are given by taking uniform samples of the Fourier transform of $r[n]$. (A rigorous treatment can be found in R. Gray's tutorial, NOW publishers 2006.)

c) Perform an empirical study as follows. Consider a bandlimited process with covariance function $r(t) = \sin(\omega_0 t)/(\omega_0 t)$. Plot the eigenvalue distribution of $\boldsymbol{R}$ corresponding to sampling frequency $1/T$ for a suitably large N and some different values of ω_0.

How many eigenvalues of $\boldsymbol{R}$ are significant and how does the answer depend on N, T and ω_0? How many degrees of freedom (approximately) does the sampled signal have? How do these observations relate to what we know from the sampling theorem?

EL8. A communication system operates over an AWGN channel with white noise (power spectral density: $N_0/2$). The system communicates one information bit (0/1) to the receiver by sending a waveform according to the following scheme.

- To communicate a "0", the transmitter sends nothing, i.e., $s(t) = 0$ for $0 \le t \le T$.
- To communicate a "1", the transmitter flips a (fair) coin. If the coin shows "head", the transmitter sends the waveform $s_a(t)$, and if the coin shows "tail," it sends the waveform $s_b(t)$ during $0 \le t \le T$.

Suppose the transmitted bit is equally likely to be a "0" or "1". Also assume that $s_a(t)$ and $s_b(t)$ have the same energy E:

$$\int_0^T s_a^2(t)dt = \int_0^T s_b^2(t) = E$$

Determine the optimal receiver and draw a block diagram. Discuss possible approximations.

EL9. Prove the formula claimed in class

$$\left(\frac{N_0}{2}\boldsymbol{I} + \boldsymbol{H}\boldsymbol{H}^T\right)^{-1} = \frac{2}{N_0}\left(\boldsymbol{I} - \boldsymbol{H}\left(\frac{N_0}{2}\boldsymbol{I} + \boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{H}^T\right) \to \Pi_{\boldsymbol{H}}^{\perp}$$

(hint: use the matrix inversion lemma).