Delft University of Technology
Faculty of Electrical Engineering, Mathematics, and Computer Science
Circuits and Systems Group

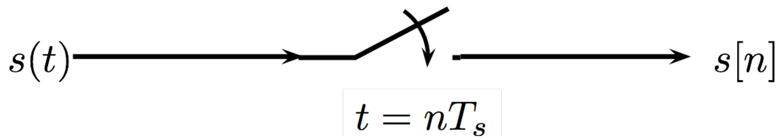
ET 4386 Estimation and Detection - ANSWERS

April 13th 2017, 13.30–16:30

This is a closed book exam. One double sided self-handwritten A4 formula sheet is allowed.

This exam has four questions (40 points in total).

Question 1 (10 points)



Continuous time sinusoids of the form $a_i \cos(2\pi f_i t) + b_i \sin(2\pi f_i t)$ are summed to produce a signal $s(t)$:

$$s(t) = \sum_{i=1}^{M} a_i \cos(2\pi f_i t) + \sum_{i=1}^{M} b_i \sin(2\pi f_i t).$$

This signal is sampled, as shown in the above picture, at the sampling instants $t = nT_s$ to get the discrete time signal

$$s[n] = \sum_{i=1}^{M} a_i \cos(2\pi f_i T_s n) + \sum_{i=1}^{M} b_i \sin(2\pi f_i T_s n).$$

Typically, the samples are taken over an interval $0 \leq t \leq NT_s$ in a "noisy" environment to produce the samples

$$x[n] = \sum_{i=1}^{M} a_i \cos(2\pi f_i T_s n) + \sum_{i=1}^{M} b_i \sin(2\pi f_i T_s n) + w[n] \quad n = 0, 1, \ldots, N-1$$

where $w[n]$ is white Gaussian noise with variance σ^2. Further, we assume the following

- The sampling period $T_s = 1$ second.
- The "known"frequencies are harmonically related as $f_i = i/N$.
- The number of sinusoids $2M < N$.

We wish to estimate the amplitudes $\boldsymbol{\theta} = [a_1, a_2, \ldots, a_M, b_1, b_2, \ldots, b_M]^T$ of these sinusoids from the noisy observations $\mathbf{x} = [x[0], x[1], \ldots, x[N-1]]^T$. This estimation problem can be written as a linear model

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}.$$

You might need the following hint to solve this question.

Hint: Use the fact that $\mathbf{H}$ matrix that you will construct is orthogonal, that is, $\mathbf{H}^T\mathbf{H} = \frac{N}{2}\mathbf{I}$

(2.5 p) (a) Set up a linear model of the form

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$$

by giving the entries of $\mathbf{H}$. Derive the least squares estimator (LSE) of $\boldsymbol{\theta}$. Also, explicitly give the expressions of the least squares estimates $\hat{a}_i$ and $\hat{b}_i$.

(2.5 p) (b) Is the LSE unbiased? Also, compute the error covariance matrix $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$.

(2.5 p) (c) Now, let us assume that the amplitudes a_i and b_i are random variables with a prior probability density function. That is, we will now assume that $\boldsymbol{\theta}$ is Gaussian distributed as

$$\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \sigma_\theta^2 \mathbf{I}).$$

Derive the minimum mean squared error (MMSE) estimator of $\boldsymbol{\theta}$. Explicitly give the expressions of the MMSE estimates of $\hat{a}_i$ and $\hat{b}_i$.

(2.5 p) (d) Compute the posterior covariance matrix $\mathbf{C}_{\boldsymbol{\theta}|\mathbf{x}}$. Compare this posterior covariance matrix with the error covariance matrix derived in *part b* of this question. What happens to the MMSE estimates as $\sigma_\theta^2 \to \infty$?

Solution

(a) To formulate the problem as a linear model, we let

$$\mathbf{H} = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ \cos\frac{2\pi}{N} & \cdots & \cos\frac{2\pi M}{N} & \sin\frac{2\pi}{N} & \cdots & \sin\frac{2\pi M}{N} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \cos\frac{2\pi(N-1)}{N} & \cdots & \cos\frac{2\pi M(N-1)}{N} & \sin\frac{2\pi(N-1)}{N} & \cdots & \cos\frac{2\pi M(N-1)}{N} \end{bmatrix},$$

where it can be shown that the columns of $\mathbf{H}$ are orthogonal. Therefore, the least-squares estimator of $\boldsymbol{\theta}$ given by

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$$

simplifies to

$$\hat{\boldsymbol{\theta}} = \frac{2}{N}\mathbf{H}^T\mathbf{x}.$$

This means that, we have

$$\hat{a}_i = \frac{2}{N}\sum_{n=0}^{N-1} x[n]\cos\frac{2\pi ik}{N}$$

and

$$\hat{b}_i = \frac{2}{N}\sum_{n=0}^{N-1} x[n]\sin\frac{2\pi ik}{N}.$$

(b) Since

$$E(\hat{a}_i) = \frac{2}{N}\mathbf{h}_i^T\mathbf{H}\boldsymbol{\theta} = a_i$$

and

$$E(\hat{b}_i) = \frac{2}{N}\mathbf{h}_i^T\mathbf{H}\boldsymbol{\theta} = b_i,$$

where $\mathbf{h}_i^T$ forms the ith row of $\mathbf{H}$, the LSE is unbiased. Also, for linear models we know that the covariance matrix

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = \sigma^2(\mathbf{H}^T\mathbf{H})^{-1}) = \frac{2\sigma^2}{N}\mathbf{I}.$$

(c) For Gaussian models with Gaussian prior, the MMSE estimate is computed by evaluating

$$\begin{aligned}\hat{\boldsymbol{\theta}} = E(\boldsymbol{\theta}|\mathbf{x}) = \sigma_\theta^2\mathbf{H}^T(\mathbf{H}\sigma_\theta^2\mathbf{H}^T + \sigma^2\mathbf{I})^{-1}\mathbf{x} &= (\mathbf{H}^T\sigma^{-2}\mathbf{H} + \sigma_\theta^{-2}\mathbf{I})^{-1}\sigma^{-2}\mathbf{H}^T\mathbf{x} \\ &= \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma_\theta^2} + \frac{N}{2\sigma^2}}\mathbf{H}^T\mathbf{x}\end{aligned}$$

and the expressions for $\hat{a}_i$ and $\hat{b}_i$ are given by

$$\hat{a}_i = \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma_\theta^2} + \frac{N}{2\sigma^2}}\sum_{n=0}^{N-1} x[n]\cos\frac{2\pi ik}{N}$$

and

$$\hat{b}_i = \frac{\frac{1}{\sigma^2}}{\frac{1}{\sigma_\theta^2} + \frac{N}{2\sigma^2}}\sum_{n=0}^{N-1} x[n]\sin\frac{2\pi ik}{N}.$$

(d) The posterior covariance matrix is given by

$$\mathbf{C}_{\boldsymbol{\theta}|\mathbf{x}} = \sigma_\theta^2\mathbf{I} - \sigma_\theta^2\mathbf{H}^T(\mathbf{H}\sigma_\theta^2\mathbf{H}^T + \sigma^2\mathbf{I})^{-1}\mathbf{H}\sigma_\theta^2 = \frac{1}{\frac{1}{\sigma_\theta^2} + \frac{N}{2\sigma^2}}\mathbf{I}.$$

As $\sigma_\theta^2 \to \infty$, there is no prior knowledge, for which the posterior covariance matrix is same as the error covariance matrix of the LSE.

Question 2 - 10 points

(2 p) (a)

$$p(\mathbf{y};\lambda) = \left(\frac{\lambda}{2}\right)^N e^{-\lambda \sum_{n=0}^{N-1} |y[n]|}$$

$$\frac{\partial \ln p(\mathbf{y};\lambda)}{\partial \lambda} = \frac{N}{\lambda} - \sum_{n=0}^{N-1} |y[n]|$$

$$\frac{\partial^2 \ln p(\mathbf{y};\lambda)}{\partial \lambda^2} = -\frac{N}{\lambda^2}$$

$$\text{var}\left(\hat{\lambda}\right) \geq \frac{1}{-E\left[\frac{\partial^2 \ln p(\mathbf{y};\lambda)}{\partial \lambda^2}\right]} = \frac{1}{I(\lambda)} = \frac{\lambda^2}{N}$$

(2 p) (b) To find an MVU that meets the CRLB ,

$$\frac{\partial \ln p(\mathbf{y};\lambda)}{\partial \lambda} = \frac{N}{\lambda} - \sum_{n=0}^{N-1} |y[n]|$$

should be put in the form

$$\frac{\partial \ln p(\mathbf{y};\lambda)}{\partial \lambda} = I(\lambda)\left(f(\mathbf{y}) - \lambda\right).$$

This is not possible for this density. There does thus not exist an unbiased estimator that can achieve the CRLB.

(2 p) (c)

$$\frac{\partial \ln p(\mathbf{y};\lambda)}{\partial \lambda} = \frac{N}{\lambda} - \sum_{n=0}^{N-1} |y[n]| = 0$$

$$\lambda_{MLE} = \frac{N}{\sum_{n=0}^{N-1} |y[n]|}$$

(2 p) (d)

$$\lambda_{MAP} = \arg\max_{\lambda} p(\lambda|\mathbf{y}) = \arg\max_{\lambda} p(\mathbf{y}|\lambda)p(\lambda)$$

The prior distribution $p(\lambda)$ influences the MAP estimator, depending on the value c.

1) $\lambda_{MAP} = \min\left[\frac{N}{\sum_{n=0}^{N-1}|y[n]|}, \beta\right]$

2) $\lambda_{MAP} = \max\left[c, \min\left[\frac{N}{\sum_{n=0}^{N-1}|y[n]|}, \beta + c\right]\right]$

(2 p) (e)

$$E[\lambda|\mathbf{y}] = \frac{\int_0^\infty \lambda p(\mathbf{y}|\lambda)p(\lambda)d\lambda}{\int_0^\infty p(\mathbf{y}|\lambda)p(\lambda)d\lambda} = \frac{\int_c^{c+\beta} \lambda p(\mathbf{y}|\lambda)d\lambda}{\int_c^{c+\beta} p(\mathbf{y}|\lambda)p(\lambda)d\lambda}$$

For $c = 0$ and $\beta \to \infty$ this is the MMSE estimator. However, for limited β ,$p(\mathbf{y}|\lambda)$ gets truncated. For $c > 0$ the prior also influences the estimate for λ. We discard part of the data. As a result, small values of λ are less likely.

Question 3 - 10 points

We would like to transmit binary information over a noisy channel. The information will be transmitted in sequences of N samples. If, in this sequence of N samples, the signal is detected (that is $\mathcal{H}_1$), a '1' is decided, and if no signal is detected (that is $\mathcal{H}_0$) a '0' is decided.

First, we test the use of the signal $s_1[n] = A[n]$ with $A[n] \sim \mathcal{N}(0, \sigma_A^2)$ and uncorrelated across time. That is, every sample n that a signal sample is transmitted, this consists of a new (uncorrelated) realisation of the random variable $A[n]$.

In vector form the hypotheses can be written as

$$\begin{aligned}\mathcal{H}_0 : &\quad \mathbf{x} = \mathbf{w}\\ \mathcal{H}_1 : &\quad \mathbf{x} = \mathbf{s}_1 + \mathbf{w},\end{aligned}$$

with $\mathbf{w} = [w[0], w[1], ..., w[N-1]]^T$, $\mathbf{s}_1 = [A[0], A[1], ..., A[N-1]]^T$, $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ and $\mathbf{s}_1 \sim \mathcal{N}(\mathbf{0}, \sigma_A^2\mathbf{I})$.

(1 p) (a) We have the following linear model: $\mathbf{x} = \mathbf{s}_1 + \mathbf{w}$ with $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ and $\mathbf{s}_1 \sim \mathcal{N}(\mathbf{0}, \sigma_A^2\mathbf{I})$. For this Gaussian linear model, the MMSE estimator is given by

$$\hat{\mathbf{s}}_1 = E[\mathbf{s}_1|\mathbf{x}] = \mathbf{C_A}(\mathbf{C_A} + \mathbf{C_W})^{-1}\mathbf{x} = \sigma_A^2\mathbf{I}(\sigma_A^2\mathbf{I} + \sigma^2\mathbf{I})^{-1}\mathbf{x} = \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2}\mathbf{x}$$

(2 p) (b) Under $\mathcal{H}_0$ the distribution of $\mathbf{x}$ is $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, and under $\mathcal{H}_1$ the distribution of $\mathbf{x}$ is $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, (\sigma_A^2 + \sigma^2)\mathbf{I})$. The NP likeilood ratio test is given by

$$\frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} \geq \gamma.$$

From here, the Neyman-Pearson detector $T(\mathbf{x})$ can be derived and is given by

$$T(\mathbf{x}) = \hat{\mathbf{s}}_1^T\mathbf{x} = \frac{\sigma_A^2}{\sigma_A^2 + \sigma^2}\mathbf{x}^T\mathbf{x}.$$

From this formulation it is clear that the optimal detector is an estimator followed by a correlator. An alternative formulation is obtained by dividing $T(\mathbf{x})$ by $frac\sigma_A^2\sigma_A^2 + \sigma^2$ leading to

$$T_2(\mathbf{x}) = \mathbf{x}^T\mathbf{x}.$$

In this formulation it is clear that under this statistical model we have a energy detector.

(2 p) (c) Under $\mathcal{H}_0$, $\frac{T(\mathbf{x})}{\sigma^2} \sim \chi_N^2$ and under $\mathcal{H}_1$, $\frac{T(\mathbf{x})}{\sigma^2+\sigma_A^2} \sim \chi_N^2$.

$$P_{fa} = Pr\left\{T(\mathbf{x}) > \gamma'; \mathcal{H}_0\right\} = Pr\left\{\frac{T(\mathbf{x})}{\sigma^2} > \frac{\gamma'}{\sigma^2}; \mathcal{H}_0\right\} = Q_{\chi_N^2}\left(\frac{\gamma'}{\sigma^2}\right)$$

γ' is thus given by $\gamma' = Q_{\chi_N^2}^{-1}\left(P_{fa}\right)\sigma^2$. The P_D is then given by

$$P_D = Pr\left\{T(\mathbf{x}) > \gamma'; \mathcal{H}_1\right\} = Q_{\chi_N^2}\left(\frac{Q_{\chi_N^2}^{-1}\left(P_{fa}\right)\sigma^2}{\sigma_A^2+\sigma^2}\right)$$

(1 p) (d) Under $\mathcal{H}_0$ the distribution of $\mathbf{x}$ is $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$, and under $\mathcal{H}_1$ the distribution of $\mathbf{x}$ is $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I})$

(1 p) (e) We now have the linear model $\mathbf{x} = A\mathbf{1} + \mathbf{w}$.

$$\hat{\mathbf{s}}_2 = E[\mathbf{s}_2|\mathbf{x}] = \sigma_A^2\mathbf{1}\mathbf{1}^T(\sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I})^{-1}\mathbf{x}.$$

(1 p) (f)

$$T(\mathbf{x}) = \hat{\mathbf{s}}_2^T\mathbf{x} = \sigma_A^2\mathbf{1}\mathbf{1}^T(\sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I})^{-1}\mathbf{x}^T\mathbf{x} \geq \gamma''.$$

(2 p) (g)

$$\mathbf{C}_{\mathbf{s_1}|\mathbf{x}} = \sigma_A^2\left(1 - \frac{\sigma_A^2}{\sigma_A^2+\sigma^2}\right)\mathbf{I}$$

and $\mathbf{C}_{\mathbf{s_2}|\mathbf{x}} = \sigma_A^2\mathbf{1}\mathbf{1}^T - \sigma_A^2\mathbf{1}\mathbf{1}^T(\sigma_A^2\mathbf{1}\mathbf{1}^T + \sigma^2\mathbf{I})^{-1}\sigma_A^2\mathbf{1}\mathbf{1}^T = \sigma_A^2\left(1 - \frac{\sigma_A^2}{\frac{\sigma^2}{N}+\sigma_A^2}\right)\mathbf{1}\mathbf{1}^T$. The Bayesian mean-squared error (Bmse) is in this cae given by the (i,i)th component of the conditional covariance matrix, i.e., $[\mathbf{C}_{\mathbf{s_1}}|\mathbf{x}]_{i,i}$ and $[\mathbf{C}_{\mathbf{s_2}}|\mathbf{x}]_{i,i}$.

$$[\mathbf{C}_{\mathbf{s_1}}|\mathbf{x}]_{1,1} = \sigma_A^2\left(1 - \frac{\sigma_A^2}{\sigma_A^2+\sigma^2}\right)$$

and

$$[\mathbf{C}_{\mathbf{s_2}}|\mathbf{x}]_{1,1} = \sigma_A^2\left(1 - \frac{\sigma_A^2}{\frac{\sigma^2}{N}+\sigma_A^2}\right).$$

As $[\mathbf{C}_{\mathbf{s_2}}|\mathbf{x}]_{1,1} \leq [\mathbf{C}_{\mathbf{s_1}}|\mathbf{x}]_{1,1}$, the second detector will always be more effective. This is due to the fact that it can exploit the correlation in signal $\mathbf{s}_2$.

Question 4 (10 points)

In an optical communication system, bits are signaled by turning on and off a laser source. Suppose during a bit period, a laser transmits n photons according a Poisson process with "probability mass function"

$$P(N = n) = e^{-\lambda}\lambda^n/n!.$$

Here, n can only be "integers" and $n!$ is read as "n factorial". We want to count the number of photons n that arrive and use this observed data to make a decision. We do this by solving the binary hypothesis testing problem

$$\mathcal{H}_0 : \lambda = \lambda_0; \quad \text{photon arrival rate when laser is off (0 bit)}$$

and

$$\mathcal{H}_1 : \lambda = \lambda_1; \quad \text{photon arrival rate when laser is on (1 bit).}$$

Assume that $\lambda_1 > \lambda_0$.

(3.5 p) (a) Find the Neyman-Pearson test to choose hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$ by showing that the number of photons in a bit period a "test statistic".

(3.5 p) (b) Give the expressions for the probability of detection P_d and probability of false alarm P_f for this hypothesis testing problem. It is sufficient if you give the expressions as a function of the threshold.

(3 p) (c) Assuming that both the hypotheses occur with equal probability, that is, $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 1/2$, derive the "test statistic". Also, give the expression for the probability of error P_e.

Solution

(a) Under the two hypotheses, we have the following Poisson distributions

$$\mathcal{H}_0 : P(N = n) = e^{-\lambda_0}\lambda_0^n/n!$$

and

$$\mathcal{H}_1 : P(N = n) = e^{-\lambda_1}\lambda_1^n/n!$$

for $n = 0, 1, 2, \ldots$. Then, based on the likelihood ratio test, we decide $\mathcal{H}_1$ if

$$L(n) = \left(\frac{\lambda_1}{\lambda_0}\right)^n e^{-(\lambda_1 - \lambda_0)} \geq \gamma$$

or, equivalently (assuming $\lambda_1 > \lambda_0$), we decide $\mathcal{H}_1$ if

$$n \geq \frac{\ln\gamma + \lambda_1 - \lambda_0}{\ln\lambda_1 - \ln\lambda_0} = \gamma'.$$

Therefore, the test statistic equals the number of photons in a bit period.

(b) Since n takes only integer values, we can rewrite the likelihood ratio test to decide $\mathcal{H}_1$ as

$$n \geq \gamma'_I, \quad \gamma'_I = 0, 1, 2, \ldots.$$

Using this, we can express the P_d as

$$P_d = e^{-\lambda_1} \sum_{n \geq \gamma'_I} \lambda_1^n/n! = 1 - e^{-\lambda_1} \sum_{n=0}^{\gamma'_I - 1} \lambda_1^n/n!$$

and, we can express P_f as

$$P_f = e^{-\lambda_0} \sum_{n < \gamma'_I} \lambda_0^n/n! = 1 - e^{-\lambda_0} \sum_{n=0}^{\gamma'_I - 1} \lambda_0^n/n!.$$

(c) The test statistic in the Bayesian setting with $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 1/2$, is

$$L(n) = \left(\frac{\lambda_1}{\lambda_0}\right)^n e^{-(\lambda_1 - \lambda_0)} \geq 1$$

or, equivalently (assuming $\lambda_1 > \lambda_0$), we decide $\mathcal{H}_1$ if

$$n \geq \frac{\lambda_1 - \lambda_0}{\ln\lambda_1 - \ln\lambda_0}.$$

The probability of error, P_e, is

$$P_e = P(\mathcal{H}_0)P_f + P(\mathcal{H}_1)(1 - P_d) = 0.5 + 0.5(P_f - P_d),$$

where the expressions for P_f and P_d are same as before.