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

ET 4386 Estimation and Detection

27 January 2017, 9:00–12:00

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

We observe the data $x[n]$, for $n = 0, .., N-1$, where $x[n]$ has the probability density function (pdf)

$$p(x[n]; \lambda) = \begin{cases} \lambda \exp(-\lambda x[n]) & x[n] > 0 \\ 0 & x[n] < 0, \end{cases}$$

with $E[x[n]] = \frac{1}{\lambda}$ and $E[x[n]^2] = \frac{2}{\lambda^2}$. The $x[n]$'s are independent. One is interested in estimating the parameter λ.

(2 p) (a) Determine the Cramér-Rao lower bound (CRLB) for λ as a function of N.

(2 p) (b) Does an unbiased estimator for λ exist that achieves the CRLB for a finite number N? If not, explain why.

(2 p) (c) Determine the maximum likelihood estimator (MLE) for λ using N data samples.

Now, we will investigate the use of a Bayesian estimator for estimating λ. We therefore assume a conditional pdf

$$p(x[n]|\lambda) = \begin{cases} \lambda \exp(-\lambda x[n]) & x[n] > 0 \\ 0, & x[n] < 0. \end{cases}$$

The $x[n]$'s are independent when conditioned on λ. The variable λ is now assumed to be a random variable with the prior distribution $p(\lambda)$ given by

$$p(\lambda) = \begin{cases} a \exp(-\lambda a) & \lambda > 0 \\ 0 & \lambda < 0, \end{cases}$$

with variance $\text{var}[\lambda] = \frac{1}{a^2}$.

(2 p) (d) Determine the MAP estimator of λ using N data samples.

(2 p) (e) Explain the relation between the MAP estimator and the above derived MLE estimator in terms of the variance $\text{var}[\lambda]$ and the number of data samples N.

Question 2 - 10 points

Consider a known signal $As[n]$ for $n = 0, 1$. At the receiver, this known signal is contaminated with noise $w[n]$. We would like to detect whether $As[n]$ is present or not. To do so, we distinguish between two hypotheses, which can be written in a vector form as

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

where $\mathbf{x} = [x[0], x[1]]^T$, $\mathbf{w} = [w[0], w[1]]^T$, and $\mathbf{s} = [s[0], s[1]]^T$. The noise has a zero-mean bivariate Gaussian distribution, i.e., $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$ with

$$\mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

(2.5 p) (a) Determine the Neyman-Pearson detector $T(\mathbf{x})$. Also specify the complete pdf of $T(\mathbf{x})$ and an expression for the corresponding optimal probability of detection as a function of P_{fa}.

Hint: The right-tail probability of the pdf $p(t)$ is given by

$$Q(t) = \int_x^\infty p(t)dt.$$

We are interested in transmitting the signal $A\mathbf{s}$ with minimum power, while still having a pre-specified detection probability P_D.

(2.5 p) (b) Assume that $\mathbf{s} = [1, 1]^T$. What is the minimum value for A that we need such that $P_D > 10P_{FA}$?

Now, we fix A to $A = 1$ and keep the total power in $\mathbf{s}$ fixed as well, i.e., $s^2[0] + s^2[1] = 2$.

(2.5 p) (c) Give the signal $\mathbf{s} = [s[0], s[1]]^T$ that maximizes the detection probability for $s^2[0] + s^2[1] = 2$.

(2.5 p) (d) Explain for a non-diagonal correlation matrix $\mathbf{C}$, how $\mathbf{s}$ should be chosen such that P_D is optimal in the Neyman-Pearson sense.

Question 3 (10 points)

Independent bivariate Gaussian samples

$$\{\mathbf{x}[0], \mathbf{x}[1], \ldots, \mathbf{x}[N-1]\}$$

are observed with $\mathbf{x}[n] = [x_0[n], x_1[n]]^T$. Each observation is a 2×1 vector that is Gaussian distributed according to probability density function

$$p(\mathbf{x}[n]; \mathbf{C}(\rho)) = \frac{1}{2\pi}(\det\mathbf{C}(\rho))^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{x}[n]^T\mathbf{C}^{-1}(\rho)\mathbf{x}[n]\right\}.$$

and

$$\mathbf{C}(\rho) = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

(2.5 p) (a) Find the Cramér-Rao lower bound (CRLB) for the correlation coefficient ρ.

Hint: You may want to use the following matrix identity:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

For the scalar parameter case θ, with one observation, in which $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(\theta))$, the Fisher Information reduces to

$$I(\theta) = \frac{1}{2}\text{tr}\left[\left(\mathbf{C}^{-1}(\theta)\frac{\partial\mathbf{C}(\theta)}{\partial\theta}\right)^2\right],$$

where $\text{tr}[\cdot]$ is defined to be the sum of the elements along the main diagonal of a square matrix.

(2 p) (b) For $N = 1$, plot the CRLB versus the correlation coefficient $-1 < \rho < 1$. Explain what happens when $\rho \to \pm 1$ and for $\rho = 0$.

(3.5 p) (c) Derive the least squares estimator (LSE) of ρ. The LSE is found by minimizing

$$J = \text{tr}\left\{[\mathbf{S} - \mathbf{C}(\rho)]^T[\mathbf{S} - \mathbf{C}(\rho)]\right\},$$

where

$$\mathbf{S} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}[n]\mathbf{x}^T[n]$$

is the 2×2 sample covariance matrix.

Hint: The identities

$$\frac{\partial}{\partial x}\text{tr}\{\mathbf{A}^T\mathbf{B}\} = \text{tr}\left\{\mathbf{A}^T\frac{\partial \mathbf{B}}{\partial x}\right\}$$

and

$$\frac{\partial}{\partial x}\text{tr}\{\mathbf{B}^T\mathbf{B}\} = 2\text{tr}\left\{\mathbf{B}^T\frac{\partial \mathbf{B}}{\partial x}\right\}$$

will be useful.

(2 p) (d) Is it possible to determine the Best Linear Unbiased Estimator (BLUE) for this problem? If yes, explain how? If not, explain why? (You need not derive BLUE for this problem.)

Question 4 (10 points)

Assume that we wish to distinguish between the hypotheses

$$\mathcal{H}_0 : \mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$$

and

$$\mathcal{H}_1 : \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$$

based on two bivariate Gaussian samples $\mathbf{x} = [x[0], x[1]]^T$ distributed according to the probability density function

$$p(\mathbf{x}; \boldsymbol{\mu}, \mathbf{C}) = \frac{1}{2\pi}(\det\mathbf{C})^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$$

with

$$\mathbf{C} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

and $\boldsymbol{\mu} = [\mu_0, \mu_1]^T$. Assume that $\mu_1 > 0$.
Hint: You may want to use the following matrix identity to solve this question:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

(3 p) (a) Given $P(\mathcal{H}_0) = P(\mathcal{H}_1) = 1/2$, find the *maximum (conditional) likelihood detector* that minimizes the probability of error, P_e, by deciding on $\mathcal{H}_1$ if

$$p(\mathbf{x}|\mathcal{H}_1) > p(\mathbf{x}|\mathcal{H}_0).$$

(3.5 p) (b) Now for $\rho = 0$, find the minimum P_e detector. Using a 2D plot of $x[0]$ vs. $x[1]$, illustrate decision regions (to decide $\mathcal{H}_1$ or $\mathcal{H}_0$). Show that the decision region boundary is a line that is the *perpendicular bisector* of the line segment from $\mathbf{0}$ to $\boldsymbol{\mu}$.

(3.5 p) (c) Now assume $\boldsymbol{\mu} = [1, 1]^T$, using a 2D plot of $x[0]$ vs. $x[1]$, illustrate decision regions (to decide $\mathcal{H}_1$ or $\mathcal{H}_0$) for two values of the correlation coefficient, i.e., for $\rho = 0.2$ and $\rho = 0.5$. Explain what happens to the decision region boundary when the value of ρ changes. Also, explain why this happens.