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

ET 4386 Estimation and Detection - Answers

Januray 25th 2019, 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).

You might need the expression for the Gaussian probability density function in some questions. Below, we therefore give the general definition:

Let $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{C})$ be an $N \times 1$ random Gaussian distributed vector with $N \times 1$ mean vector $\mathbf{m}$ and $N \times N$ covariance matrix $\mathbf{C}$. The probability density function of $\mathbf{x}$ is then given by

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} (\det \mathbf{C})^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T \mathbf{C}^{-1}(\mathbf{x}-\mathbf{m})\right]$$

Question 1 (10 points)

Let

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

with $\mathbf{w} \sim N(\mathbf{0}, \mathbf{C})$ and $\mathbf{s} \sim N(\boldsymbol{\mu}_s, \eta\mathbf{C})$ with η a scalar with $\eta \geq 0$, $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{s} \in \mathbb{R}^N$ and $\mathbf{w} \in \mathbb{R}^N$.

(2 pts) (a) The likelihood ratio is given by

$$\frac{\frac{1}{(2\pi)^{N/2} det^{1/2}(\mathbf{C}(1+\eta))} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_s)^T(\mathbf{C}(1+\eta))^{-1}(\mathbf{x}-\boldsymbol{\mu}_s)\right]}{\frac{1}{(2\pi)^{N/2} det^{1/2}(\mathbf{C})} \exp\left[-\frac{1}{2}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}\right]} > \lambda$$

Taking the logarithm and only maintaining the data dependent terms, we get

$$-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_s)^T(\mathbf{C}(1+\eta))^{-1}(\mathbf{x}-\boldsymbol{\mu}_s) + \frac{1}{2}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} > \lambda'$$

Finally resulting in (taking only the data-dependent terms)

$$T(\mathbf{x}) = \underbrace{\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}\frac{\frac{1}{2}\eta}{1+\eta}}_{\text{quadratic in } \mathbf{x}} + \underbrace{\frac{1}{1+\eta}\mathbf{x}^T C^{-1}\boldsymbol{\mu}_s}_{\text{linear in } \mathbf{x}} > \lambda''$$

(0.5 pts) (b) We can write $\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} = \mathbf{y}^T\mathbf{y}$ with $\mathbf{y} = \mathbf{C}^{-1/2}\mathbf{x}$. Calculating $E[\mathbf{y}\mathbf{y}^T] = E[\mathbf{C}^{-1/2}\mathbf{x}\mathbf{x}^T\mathbf{C}^{-1/2}] = \mathbf{I}$. This means that the elements in the vector $\mathbf{y}$ are completely uncorrelated (and independent as $\mathbf{x}$ is Gaussian), and, all elements have variance 1. The product $\mathbf{y}^T\mathbf{y} = \sum_{n=1}^{N} y^2[n]$ is thus a sum of squared zero-mean Gaussian samples with variance 1 and is this χ^2 distributed.

(0.5 pts) (c) Using the same argumentation as in question b, we can write $\mathbf{y} = \mathbf{C}^{-1/2}\mathbf{x}$. However, now we get $E[\mathbf{y}\mathbf{y}^T] = E[\mathbf{C}^{-1/2}\mathbf{x}\mathbf{x}^T\mathbf{C}^{-1/2}] = \mathbf{I}(\eta+1)$. The product $\mathbf{y}^T\mathbf{y} = \sum_{n=1}^{N} y^2[n]$ is now thus a sum of squared zero-mean Gaussian samples with variance $(1+\eta)$ and is this χ^2 distributed if we normalise the variances to one and divide the product by $(1+\eta)$: $\frac{\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}}{\eta+1} \sim \chi^2_N$.

(2 pts) (d) From question a) we get

$$T(\mathbf{x}) = \mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}\frac{\frac{1}{2}\eta}{1+\eta} > \lambda''.$$

which is equivalent to

$$T(\mathbf{x}) = \mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} > \lambda'''.$$

by neglecting the constant.

For P_{FA} it holds that (using the given statement from Questions b) and c)) $P_{FA} = P(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} \geq \lambda'''; \mathcal{H}_0) = Q_{\chi^2_N}(\lambda''')$. And thus we get $\lambda''' = Q^{-1}_{\chi^2_N}(P_{FA})$.

For P_D we then get $P_D = P(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x} \geq \lambda'''; \mathcal{H}_1) = Q_{\chi^2_N}(\frac{\lambda'''}{1+\eta}) = Q_{\chi^2_N}\left(\frac{Q^{-1}_{\chi^2_N}(P_{FA})}{1+\eta}\right)$

(2 pts) (e) From question a) we get $T(\mathbf{x}) = \mathbf{x}^T\mathbf{C}^{-1}\boldsymbol{\mu}_s > \lambda''$

$E[T(\mathbf{x}); \mathcal{H}_0] = E[\mathbf{w}^T\mathbf{C}^{-1}\boldsymbol{\mu}_s] = E[\mathbf{w}^T]\mathbf{C}^{-1}\boldsymbol{\mu}_s = 0.$

$E[T(\mathbf{x}); \mathcal{H}_1] = E[\mathbf{x}^T\mathbf{C}^{-1}\boldsymbol{\mu}_s] = E[(\mathbf{w}+\boldsymbol{\mu}_s)^T\mathbf{C}^{-1}\boldsymbol{\mu}_s] = E[\mathbf{w}^T]\mathbf{C}^{-1}\boldsymbol{\mu}_s + \boldsymbol{\mu}_s^T C^{-1}\boldsymbol{\mu}_s = \boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s.$

$var[T(\mathbf{x}); \mathcal{H}_0] = E[\boldsymbol{\mu}_s\mathbf{C}^{-1}\mathbf{w}\mathbf{w}^T\mathbf{C}^{-1}\boldsymbol{\mu}_s] = \boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s$

$var[T(\mathbf{x}); \mathcal{H}_1] = E[\boldsymbol{\mu}_s\mathbf{C}^{-1}(\mathbf{w}+\boldsymbol{\mu}_s)(\mathbf{w}+\boldsymbol{\mu}_s)^T\mathbf{C}^{-1}\boldsymbol{\mu}_s] - E[T(\mathbf{x}); \mathcal{H}_1] = \boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s.$

(1 pts) (f) $P_{FA} = P(T(\mathbf{x}) \geq \lambda'''; \mathcal{H}_0) = Q\left(\frac{\lambda'''}{\sqrt{\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}}\right)$. Thus $\lambda''' = Q^{-1}(P_{FA})\sqrt{\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}$.

$P_D = P(T(\mathbf{x}) \geq \lambda'''; \mathcal{H}_1) = Q\left(\frac{\lambda'''-\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}{\sqrt{\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}}\right) = Q\left(\frac{Q^{-1}(P_{FA})\sqrt{\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}-\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}{\sqrt{\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}}\right) = Q\left(Q^{-1}(P_{FA}) - \sqrt{\boldsymbol{\mu}_s^T\mathbf{C}^{-1}\boldsymbol{\mu}_s}\right)$

(2 pts) (g) The vector $\boldsymbol{\mu}_s$ that leads to the best detection performance, is the eigenvector from $\mathbf{C}$ that corresponds to the smallest eigenvale of $\mathbf{C}$. Calculating $\mathbf{C}\mathbf{v}_1 = (1+\rho)\mathbf{v}_1$ and $\mathbf{C}\mathbf{v}_2 = (1-\rho)\mathbf{v}_2$, we learn that $\mathbf{v}_1$ has eigenvalue $1+\rho$ and $\mathbf{v}_2$ has eigenvalue $1-\rho$. The vector $\boldsymbol{\mu}_s$ that leads to the best detection performance is thus $\mathbf{v}_2$.

Question 2 (8 points)

(2 pts) (a) $\frac{a_1 x e^{-a_1 x^2/2}}{a_0 x e^{-a_0 x^2/2}} \geq \lambda$.

$x^2(a_0 - a_1) \geq \lambda'$

$T(x) = x^2 \geq \lambda''$

(2 pts) (b) $P_{fa} = P(x^2 \geq \lambda''; \mathcal{H}_0) = P(x \geq \lambda''; \mathcal{H}_0) = \int_{\lambda''}^{\infty} a_0 x e^{-a_0 x^2/2} dx = \left[-e^{-a_0 x^2/2}\right]_{\lambda''}^{\infty} = e^{-a_0 \frac{(\lambda'')^2}{2}} = P_{fa}$.

$\lambda'' = \left(\frac{2}{a_0} \log\left(1/P_{FA}\right)\right)^{1/2}$

(2 pts) (c) $P_D = P(x^2 \geq \lambda''; \mathcal{H}_1) = P(x \geq \lambda''; \mathcal{H}_1) = \int_{\lambda''}^{\infty} a_1 x e^{-a_1 x^2/2} dx = e^{-a_1 \frac{(\lambda'')^2}{2}} = e^{\frac{a_1}{a_0} \log(P_{FA})}$

Question 3 (10 points)

In this question we revisit the problem of finding the DC component from a signal corrupted by white Gaussian noise when the noise variance is *unknown.*

Given N data samples drawn i.i.d. from a $\mathcal{N}(\mu, \sigma^2)$ distribution, where both μ and σ^2 are unknown, determine the following:

(1 pt) (a) Write down the data model for $\mathbf{x} = \left[x[0], \ldots, x[N-1]\right]^T$.

Solution:

$$\mathbf{x} = \mu\mathbf{1} + \mathbf{w}, \qquad \text{with} \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}),$$

or,

$$x[n] = \mu + w[n], \qquad \text{with} \quad w[n] \sim \mathcal{N}(0, \sigma^2), \quad \text{for} \quad n = 0, \ldots, N-1.$$

(1 pt) (b) Show that the sample mean

$$\bar{x} = \left(\frac{1}{N}\right)\mathbf{x}^T\mathbf{1} = \left(\frac{1}{N}\right)\sum_{n=0}^{N-1} x[n],$$

where $\mathbf{1} = [1, \ldots, 1]^T$, is an unbiased estimator for μ.

Solution:

$$E[\bar{x}] = \left(\frac{1}{N}\right)E[\mathbf{x}^T\mathbf{1}] = \left(\frac{1}{N}\right)\sum_{n=0}^{N-1} E\left[x[n]\right], = \mu, \tag{1}$$

since $x[n] \sim \mathcal{N}(\mu, \sigma^2)$, for all n.

In the following call

$$V = (\mathbf{x} - \bar{x})^T(\mathbf{x} - \bar{x}) = \sum_{n=0}^{N-1}(x[n] - \bar{x})^2.$$

(3 pts) (c) Show that the estimator

$$\hat{\sigma}^2_{(N)} = \left(\frac{1}{N}\right)V,$$

is a *biased* estimator for σ^2.

Solution: The question can be rephrased as to calculate $E[\hat{\sigma}^2_{(N)}] = (1/N)E[V]$, to verify if it is equal to σ^2. Let us rewrite V as:

$$\begin{aligned} V &= ((\mathbf{x} - \mu) - (\bar{x} - \mu))^T ((\mathbf{x} - \mu) - (\bar{x} - \mu)) \\ &= \sum_{n=0}^{N-1} ((x[n] - \mu) - (\bar{x} - \mu))^2 . \end{aligned}$$

Then we have:

$$E[V] = \sum_{n=0}^{N-1} E\Big[(x[n]-\mu)^2\Big] - 2\sum_{n=0}^{N-1} E\Big[(x[n]-\mu)(\bar{x}-\mu)\Big] + \sum_{n=0}^{N-1} E\Big[(\bar{x}-\mu)^2\Big] \tag{2}$$

We identify the last term in (2) as $N\text{var}[\bar{x}]$. Then we calculate:

$$\text{var}[\bar{x}] = \frac{1}{N^2}\text{var}[\mathbf{x}^T\mathbf{1}] = \frac{1}{N^2}\sum_{n=0}^{N-1}\text{var}[x_n] = \frac{\sigma^2}{N},$$

which holds by the independence of x_n, for $n = 0, \ldots, N-1$. For the second term in (2) we first calculate:

$$\begin{aligned} E\Big[(x[n] - \mu)(\bar{x} - \mu)\Big] &= E\left[(x[n] - \mu)\left(\frac{1}{N}\sum_{m=0}^{N-1}(x[m] - \mu)\right)\right] \\ &= \frac{1}{N}E\Big[(x[n] - \mu)^2\Big] \\ &\quad + \frac{1}{N}\sum_{m\neq n} E\Big[(x[n] - \mu)(x[m] - \mu)\Big] \\ &= \frac{\sigma^2}{N}, \end{aligned}$$

since by independence $x[n]$ and $x[m]$ are uncorrelated for all n, m and therefore the sum over the cross terms is 0. The first term in (2) is (per definition of the variance) just $N\sigma^2$. With these results we rewrite (2) as:

$$\begin{aligned} E[V] &= N\sigma^2 - 2N\frac{\sigma^2}{N} + N\frac{\sigma^2}{N} \\ &= (N-1)\sigma^2. \end{aligned} \tag{3}$$

And thus

$$E\Big[\hat{\sigma}^2_{(N)}\Big] = \frac{1}{N}E[V] = \frac{N-1}{N}\sigma^2,$$

is therefore a *biased* estimator of σ^2.

(3 pts) (d) Show that the estimator

$$\hat{\sigma}^2_{(N-1)} = \left(\frac{1}{N-1}\right) V,$$

is an *unbiased* estimator.

Solution: By the result (3), of the previous excercise we have:

$$\begin{aligned} E\left[\hat{\sigma}^2_{(N-1)}\right] &= \frac{1}{N-1} E[V] \\ &= \frac{N-1}{N-1}\sigma^2 \\ &= \sigma^2, \end{aligned}$$

and $\hat{\sigma}^2_{(N-1)}$ is therefore an *unbiased* estimator of σ^2.

(2 pts) (e) Show that the estimator $\hat{\sigma}^2_{(N)}$ attains lower mean squared error (MSE), as an estimator of the variance σ^2, than the estimator $\hat{\sigma}^2_{(N-1)}$. *Hint, consider that* $E[V^2] = (N^2 - 1)\sigma^4$.

Solution: Using the hint we can calculate the mean squared error of the two estimators. The MSE of $\hat{\sigma}^2_{(N)} = 1/(N)V$ is:

$$\begin{aligned} E\left[(\hat{\sigma}^2_{(N)} - \sigma^2)^2\right] &= E\left[\frac{V^2}{(N)^2} - 2\frac{V}{N}\sigma^2 + \sigma^4\right] \\ &= \frac{(N^2-1)\sigma^4}{(N)^2} - 2\frac{(N-1)\sigma^2}{N}\sigma^2 + \sigma^4 \\ &= \frac{2N-1}{N^2}\sigma^4. \end{aligned}$$

The MSE of $\hat{\sigma}^2_{(N-1)} = 1/(N-1)V$ is:

$$\begin{aligned} E\left[(\hat{\sigma}^2_{(N-1)} - \sigma^2)^2\right] &= E\left[\frac{V^2}{(N-1)^2} - 2\frac{V}{N-1}\sigma^2 + \sigma^4\right] \\ &= \frac{(N^2-1)\sigma^4}{(N-1)^2} - 2\frac{(N-1)\sigma^2}{N-1}\sigma^2 + \sigma^4 \\ &= \frac{2}{N-1}\sigma^4. \end{aligned}$$

For all positive N we have $2(N-1)/N^2 < 2/(N-1)$ and we conclude that $\hat{\sigma}^2_{(N)}$ has lower MSE than $\hat{\sigma}^2_{(N-1)}$, even tough the former is a *biased* estimator of the variance, and the later is an *unbiased* estimator of the variance.

Question 4 (10 points)

Solution a

Let $\mathbf{y} = \mathbf{x} - \alpha\mathbf{s} + \mathbf{w} \sim \mathcal{N}(\mathbf{x} - \alpha\mathbf{s}, \mathbf{C})$, then

$$p\left(\mathbf{y};\alpha\right) = \frac{1}{\left(2\pi\right)^{\frac{N}{2}}\left(\det\mathbf{C}\right)^{\frac{1}{2}}}\exp\left[-\frac{1}{2}\left(\mathbf{x}-\alpha\mathbf{s}\right)^T\mathbf{C}^{-1}\left(\mathbf{x}-\alpha\mathbf{s}\right)\right]$$

The CRLB is given by

$$\text{var}(\hat{\alpha}) \geq \frac{1}{\mathbb{E}[\frac{\partial^2 \ln\, p(\mathbf{y};\alpha)}{\partial\alpha^2}]} \tag{4}$$

Taking the first derivative, we have

$$\begin{aligned}\frac{\partial \ln\, p(\mathbf{y};\alpha)}{\partial\alpha} &= -0.5\frac{\partial \ln\,(\det\,\mathbf{C})}{\partial\alpha} - 0.5\frac{\partial}{\partial\alpha}\left[(\mathbf{x}-\alpha\mathbf{s})^T\mathbf{C}^{-1}(\mathbf{x}-\alpha\mathbf{s})\right] \\ &= -0.5\frac{\partial}{\partial\alpha}\left[(\mathbf{x}-\alpha\mathbf{s})^T\mathbf{C}^{-1}(\mathbf{x}-\alpha\mathbf{s})\right] \\ &= -0.5\frac{\partial}{\partial\alpha}\left[\mathbf{x}\mathbf{C}^{-1}\mathbf{x} + \alpha^2\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s} - 2\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x})\right] \\ &= \left[\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x} - \alpha\mathbf{s}^T\mathbf{C}\mathbf{s}\right] \end{aligned} \tag{5}$$

and differentiating again,

$$\frac{\partial^2 p(\mathbf{y};\alpha)}{\partial\alpha^2} = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$$

and subsequently, from (4), we have the lowest achievable variance as

$$\text{var}(\hat{\alpha}) \geq \left(\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}\right)^{-1} \tag{6}$$

Solution b

Observe that (5) can be rewritten in the form

$$\begin{aligned}\frac{\partial \ln\, p(\mathbf{y};\alpha)}{\partial\alpha} &= \left(\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}\right)\left[\left(\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}\right)^{-1}\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x} - \alpha\right] \\ &= \mathbf{I}(\alpha)\left(\mathbf{g}(\mathbf{x}) - \alpha\right)\end{aligned}$$

where

$$\begin{aligned}\mathbf{I}(\alpha) &= \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s} \\ \mathbf{g}(\mathbf{x}) &= \left(\mathbf{s}^T\mathbf{C}\mathbf{s}\right)^{-1}\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x}\end{aligned}$$

Here, $\hat{\alpha}_{MVU} = \mathbf{g}(\mathbf{x})$ is the MVU estimator which achieves the minimum variance $\mathbf{I}^{-1}(\alpha) = (\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s})^{-1}$ and is the CRLB for the problem, which was derived in (6).

Solution c

When $\rho = 0$, the covariance $\mathbf{C}$ takes the form

$$\mathbf{C} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} = \sigma^2\mathbf{I}$$

and subsequently, the MVU estimator simplifies to the LS estimator *i.e.,*

$$\begin{aligned} \hat{\alpha}_{MVU} &= (\sigma^{-2}\mathbf{s}^T\mathbf{s})^{-1}\sigma^{-2}\mathbf{s}^T\mathbf{x} \\ &= (\mathbf{s}^T\mathbf{s})^{-1}\mathbf{s}^T\mathbf{x} \qquad (7) \\ &= \hat{\alpha}_{LS} \end{aligned}$$

Solution d

When $\rho = \sigma^2$, the MVUE is not feasible since $\mathbf{C} = \sigma^2[\mathbf{1_2}, \mathbf{1_2}]$ is not invertible