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

ET 4386 Estimation and Detection

January 24th 2020,

This exam has three questions (32 points in total) and is a closed book exam. One double sided self-handwritten A4 formula sheet is allowed.

The following formulas might be useful for some of the exercises:

- The general expression for the Gaussian probability density function:

  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

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

- The Woodbury identity may also be useful, which is as follows

$$(\mathbf{A}+\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}(\mathbf{A}^{-1}+\mathbf{C}^{-1})^{-1}\mathbf{A}^{-1},$$

  for matrices $\mathbf{A}, \mathbf{C}$ of appropriate sizes.

Answer 1 (12 points)

(Solution a) The MLE of μ is given by:

$$\hat{\mu} = \max_{\mu} p(\mathbf{x}; \mu, \mathcal{H}_1),$$
$$\hat{\mu} = \max_{\mu} \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x} - \mu\mathbf{1})^T(\mathbf{x} - \mu\mathbf{1})\right), \qquad (1)$$

where $\mathbf{x} = [x[0], \ldots, x[N-1]]^T$, is the data vector formed from the N available data samples, and we have dropped the k, l indexes for simplicity. The solution is given by taking the logarithm, set the derivative with respect to μ to 0, and solve for μ. Taking the logarithm of (1) we obtain:

$$\begin{aligned}
&\log\left\{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x} - \mu\mathbf{1})^T(\mathbf{x} - \mu\mathbf{1})\right)\right\} \\
&\qquad = -\frac{N}{2}\log\left\{2\pi\sigma^2\right\} - \frac{1}{2\sigma^2}(\mathbf{x} - \mu\mathbf{1})^T(\mathbf{x} - \mu\mathbf{1}) \\
&\qquad = -\frac{N}{2}\log\left\{2\pi\sigma^2\right\} - \frac{1}{2\sigma^2}\left(\mathbf{x}^T\mathbf{x} - 2\mu\mathbf{x}^T\mathbf{1} + \mu^2\mathbf{1}^T\mathbf{1}\right) \\
&\qquad = -\frac{N}{2}\log\left\{2\pi\sigma^2\right\} - \frac{1}{2\sigma^2}\left(\mathbf{x}^T\mathbf{x} - 2\mu N\overline{x} + \mu^2 N\right) \qquad (2)
\end{aligned}$$

where we define $\overline{x}$ as the mean of the data samples,

$$\overline{x} = \frac{1}{N}\sum_{n=0}^{N-1} x[n] = \frac{\mathbf{x}^T\mathbf{1}}{N}. \qquad (3)$$

The derivative of (2) with respect to μ is:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\mu}\left(-\frac{N}{2}\log\left\{2\pi\sigma^2\right\} - \frac{1}{2\sigma^2}\left(\mathbf{x}^T\mathbf{x} - 2\mu N\overline{x} + \mu^2 N\right)\right) \\
&\qquad = -\frac{1}{2\sigma^2}\left(-2N\overline{x} + 2N\mu\right) \qquad (4)
\end{aligned}$$

Setting (4) to 0 and solving for μ we obtain:

$$\begin{aligned}
-\frac{1}{2\sigma^2}\left(-2N\overline{x} + 2N\mu\right) &= 0 \\
\hat{\mu} &= \overline{x} \qquad (5)
\end{aligned}$$

(Solution b) The GRLT detector is obtained by replacing the unknown parameter with its maximum likelihood estimation (MLE) in the likelihood ratio:

$$L_G(\mathbf{x}) = \frac{p(\mathbf{x}; \hat{\mu}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\mu}, \mathcal{H}_0)}. \tag{6}$$

The GRLT becomes:

$$\begin{aligned} L_G(\mathbf{x}) &= \frac{p(\mathbf{x}; \hat{\mu}, \mathcal{H}_1)}{p(\mathbf{x}; \hat{\mu}, \mathcal{H}_0)} \geq \lambda \\ L_G(\mathbf{x}) &= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x} - \overline{x}\mathbf{1})^T(\mathbf{x} - \overline{x}\mathbf{1})\right)}{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{x}\right)} \geq \lambda. \end{aligned} \tag{7}$$

Taking the log of the likelihood ratio and simplifying we obtain:

$$\begin{aligned} \log\{L_G(\mathbf{x})\} &= -\frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{1}\overline{x})^T(\mathbf{x} - \mathbf{1}\overline{x}) + \frac{1}{2\sigma^2}(\mathbf{x}^T\mathbf{x}) \geq \lambda \\ &= \frac{1}{2\sigma^2}\left(\mathbf{x}^T\mathbf{x} - (\mathbf{x}^T\mathbf{x} - 2\overline{x}\mathbf{1}^T\mathbf{x} + \overline{x}^2\mathbf{1}^T\mathbf{1})\right) \geq \lambda \\ &= \frac{1}{2\sigma^2}\left(2\overline{x}N\overline{x} - N\overline{x}^2\right) = \frac{N\overline{x}^2}{2\sigma^2} \geq \lambda, \end{aligned} \tag{8}$$

which is equivalent as to make the test statistic function and resulting test to be:

$$T_G(\mathbf{x}) = |\overline{x}| \geq \lambda', \tag{9}$$

(Solution c) The likelihood ratio test is:

$$L(\mathbf{x}) = \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x} - \mu\mathbf{1})^T(\mathbf{x} - \mu\mathbf{1})\right)}{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x})^T(\mathbf{x})\right)} \geq \gamma. \tag{10}$$

Taking the log and simplifying:

$$\begin{aligned} \log\{L(\mathbf{x})\} &= \frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{1}\mu)^T(\mathbf{x} - \mathbf{1}\mu) + \frac{1}{2\sigma^2}(\mathbf{x}^T\mathbf{x}) \geq \gamma \\ &= \frac{1}{2\sigma^2}\left(\mathbf{x}^T\mathbf{x} - (\mathbf{x}^T\mathbf{x} - 2\mu\mathbf{1}^T\mathbf{x} + \mu^2\mathbf{1}^T\mathbf{1})\right) \geq \gamma \\ &= \frac{1}{2\sigma^2}\left(2\mu N\overline{x} - N\mu^2\right) \geq \gamma. \end{aligned} \tag{11}$$

To derive a test statistic function of the data we collect all terms independent of the data on the threshold side:

$$T_N(\mathbf{x}) = \overline{x} \geq \frac{2\sigma^2\gamma + N\mu^2}{2N\mu}. \tag{12}$$

Making a change of variable we can define the following equivalent threshold and test,

$$T_N(\mathbf{x}) = \overline{x} \leq \gamma'. \tag{13}$$

The inequality gets reversed because $\mu < 0$. The exact threshold does in fact not depend on μ, it gets determined in terms of the probability of false alarm P_{FA}. We see the resulting test statistic in this case does not depend on the unknown parameter μ.

(Solution d) The threshold of the NP test in (13) is derived. First,

$$P_{FA} = Pr\{\overline{x} \leq \gamma'; \mathcal{H}_0\} = 1 - Q\left(\frac{\gamma'}{\sqrt{\sigma^2/N}}\right) = Q\left(\frac{-\gamma'}{\sqrt{\sigma^2/N}}\right) \tag{14}$$

then,

$$\gamma' = -\sqrt{\frac{\sigma^2}{N}}Q^{-1}(P_{FA}). \tag{15}$$

(Solution e) We know already that the test statistic resulting from the NP solution (13) does not depend on the unknown parameter μ. The threshold does not depend on the unknown parameter either as we see from (15). The NP detector is realizable and by the NP lemma it is guaranteed to be the one that maximizes the probability of detection for any value of the parameter. It is therefore optimal. We conclude that the NP detector given by (13) with (15) and is the UMP test, namely

$$T_U(\mathbf{x}) = \overline{x} < \gamma', \tag{16}$$

(Solution f) The detection probability of the GLRT detector will always be equal or worse compared to the NP. In this case, the GLRT will be words as it evaluates $|\bar{x}|$ instead of $\bar{x}$ as the NP does. Taking $|\bar{x}|$, maps all the negative noise realizataions to the corresponding positive values. This means one does not only need to handle the negative values of the noise when comparing t the threshold, but also the positive. The total probability mass of the noise is therefoew twice as big as in the case of the NP.

Answer 2 (10 points)

(Solution a) Let $\mathbf{s} = A\mathbf{h}$, where $\mathbf{s}, \mathbf{h}$ are N dimensional vectors, and $A \sim \mathcal{N}(0, \sigma_A^2)$, then $E(\mathbf{s}) = \mathbf{0}$ and $cov(\mathbf{s}) = \mathbf{C}_s = \sigma_A^2 \mathbf{h}\mathbf{h}^T$. The signal of interest $\mathbf{s} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_s)$, is rank-one i.e., the covariance matrix $\mathbf{C}_s$ is rank-one.

(Solution b) Given the measurements $\mathbf{x} = \mathbf{s} + \mathbf{w}$, where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_w)$, where $\mathbf{x}$ and $\mathbf{s}$ are jointly Gaussian, an estimator which minimizes the Bmse is given by $\hat{\mathbf{s}} = \mathbf{C}_{sx}\mathbf{C}_{xx}^{-1}\mathbf{x}$, where

$$\begin{aligned}
\mathbf{C}_{sx} &= \mathbb{E}[\mathbf{s}(\mathbf{s}+\mathbf{w})^T] = \mathbf{C}_s \\
\mathbf{C}_{xx} &= \mathbb{E}[(\mathbf{s}+\mathbf{w})(\mathbf{s}+\mathbf{w})^T] = (\mathbf{C}_s + \mathbf{C}_w)
\end{aligned}$$

and substituting for the respective covariances, the MMSE estimator is

$$\hat{\mathbf{s}} = \sigma_A^2 \mathbf{h}\mathbf{h}^T (\sigma_A^2 \mathbf{h}\mathbf{h}^T + \sigma_w^2 \mathbf{I})^{-1} \mathbf{x} \tag{17}$$

(Solution c) Let $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the system failure and functionality, respectively, then we have

$$\begin{aligned}
\mathcal{H}_0 :&\quad \mathbf{x} = \mathbf{w} &&\text{(sensor failure)} \\
\mathcal{H}_1 :&\quad \mathbf{x} = \mathbf{s} + \mathbf{w} &&\text{(sensor functioning)}
\end{aligned}$$

(Solution d) From (c), we have

$$\begin{aligned}
\mathbf{x} &\sim \mathcal{N}(\mathbf{0}, \mathbf{C}_w) &&\text{(sensor failure)} \\
\mathbf{x} &\sim \mathcal{N}(\mathbf{0}, \sigma_A^2 \mathbf{I} + \mathbf{C}_w) &&\text{(sensor functioning)}
\end{aligned}$$

An NP detector decides $\mathcal{H}_1$ if

$$L(x) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \gamma$$

for a given threshold γ. Substituting for the expression for the Normal pdf, and taking logarithms on both sides, we have

$$\begin{aligned}
\mathbf{T}(\mathbf{x}) &= -0.5\mathbf{x}^T \Big[(\mathbf{C}_s + \mathbf{C}_w)^{-1} - \mathbf{C}_w^{-1} \Big] > \gamma \\
&= \mathbf{x}^T \Big[\mathbf{C}_w^{-1} - (\mathbf{C}_s + \mathbf{C}_w)^{-1} \Big] > \gamma' \qquad (18)
\end{aligned}$$

Using the Woodbury identity

$$(\mathbf{C}_w + \mathbf{C}_s)^{-1} = \Big[\mathbf{C}_w^{-1} - \mathbf{C}_w^{-1} (\mathbf{C}_w^{-1} + \mathbf{C}_s^{-1})^{-1} \mathbf{C}_w^{-1} \Big] \tag{19}$$

and thus (18) simplifies to

$$\mathbf{T}(\mathbf{x}) = \mathbf{x}^T \Big[\mathbf{C}_w^{-1}(\mathbf{C}_w^{-1} + \mathbf{C}_s^{-1})^{-1}\mathbf{C}_w^{-1}\Big]\mathbf{x} \ > \gamma' \tag{20}$$

Let $\hat{\mathbf{s}} = (\mathbf{C}_w^{-1} + \mathbf{C}_s^{-1})^{-1}\mathbf{C}_w^{-1}\mathbf{x}$, then

$$\begin{aligned} \hat{\mathbf{s}} &= \Big[\mathbf{C}_w^{-1} + \mathbf{C}_s^{-1}\Big]^{-1}\mathbf{C}_w^{-1}\mathbf{x} \\ &= \Big[\mathbf{C}_w^{-1}(\mathbf{C}_s + \mathbf{C}_w)\mathbf{C}_s^{-1}\Big]^{-1}\mathbf{C}_w^{-1}\mathbf{x} \\ &= \mathbf{C}_s(\mathbf{C}_s + \mathbf{C}_w)^{-1}\mathbf{x} \end{aligned} \tag{21}$$

Substituting the expression for $\hat{\mathbf{s}}$ in (18), the NP detector decides on $\mathcal{H}_1$, if the

$$\mathbf{T}(\mathbf{x}) = \mathbf{x}^T\mathbf{C}_w^{-1}\hat{\mathbf{s}} > \gamma', \tag{22}$$

where $\hat{\mathbf{s}}$ is given in (21).

(Solution e) The estimated signal is a LMMSE estimator as derived in (b), and in essence a Wiener filter. Let $\mathbf{D}$ be such that $\mathbf{C}_w^{-1} = \mathbf{D}^T\mathbf{D}$, then both $\mathbf{x}$ and $\hat{\mathbf{s}}$ are pre-whitened using $\mathbf{D}$. Subsequently, the NP detector correlates the pre-whitened received data $\mathbf{D}\mathbf{x}$ and an the pre-whitened signal estimate $\mathbf{D}\hat{\mathbf{s}}$, and compares this output against the threshold. Such an NP detector is hence termed an estimator-correlator.

(Solution f) Under the assumption $\mathbf{C}_w = \sigma_w^2\mathbf{I}$, we have

$$\begin{aligned} \hat{\mathbf{s}} &= \mathbf{C}_s(\mathbf{C}_s + \mathbf{C}_w)^{-1}\mathbf{x} \\ &= \sigma_A^2\mathbf{h}\mathbf{h}^T(\sigma_A^2\mathbf{h}\mathbf{h}^T + \sigma_w^2\mathbf{I})^{-1}\mathbf{x} \end{aligned}$$

and assuming $\mathbf{h} = [1, \mathbf{0}_{N-1}^T]^T$, we have

$$\begin{aligned} &\hat{s}[0] = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_w^2}x[0], \\ &\hat{s}[n] = 0 \qquad\qquad n = 1, 2, \ldots, N-1 \end{aligned}$$

The test statistic $\mathbf{T}(\mathbf{x})$ in (20) then simplifies to

$$\begin{aligned} \mathbf{T}(\mathbf{x}) &= \frac{1}{\sigma_w^2}\sum_{n=0}^{N-1} x[n]s[n] > \gamma' \\ &= \frac{1}{\sigma_w^2}\frac{\sigma_A^2}{\sigma_A^2 + \sigma_w^2}x^2[0] > \gamma'. \end{aligned} \tag{23}$$

The above NP detector is a zero-memory filter, that scales on the first (squared) measurement by a factor $\frac{\sigma_A^2}{\sigma_A^2 + \sigma_w^2}$, which is cummulatively compared against the threshold γ'. As $\sigma_A^2 >> \sigma_w^2$, this scaling factor equals one, and subsequently we have an energy detector based on the first sample of the measurement given by

$$\mathbf{T}(\mathbf{x}) = \frac{x^2[0]}{\sigma_w^2} > \gamma'. \tag{24}$$

Answer 3 (10 points)

(Solution a) Let $\mathbf{s} = [\sin(2\pi f0+\phi), \sin(2\pi f1+\phi)..., \sin(2\pi f(N-1)+\phi)]^T$
Leaving out the constant terms we get $\frac{\partial \ln f(\mathbf{x};A,f,\phi)}{\partial A} = \frac{\partial -\frac{1}{2}(\mathbf{x}-A\mathbf{s})^T\mathbf{C}^{-1}(\mathbf{x}-A\mathbf{s})}{\partial A} = \frac{1}{2}(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{s}+\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x}-2A\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s})$, from which it follows that $E\left[\frac{\partial \ln f(\mathbf{x};A,f,\phi)}{\partial A}\right] = 0$ and thus is the regularity conditions satisfied.

(Solution b) $\frac{\partial^2 \ln f(\mathbf{x};A,f,\phi)}{\partial A^2} = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$. The Fisher information is then given by $I(A) = E\left[\frac{\partial^2 \ln f(\mathbf{x};A,f,\phi)}{\partial A^2}\right] = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$. the CRLB is then given by $var(\hat{A}) \geq 1/I(A) = \frac{1}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}$.

(Solution c) Put $\frac{\partial \ln f(\mathbf{x};A,f,\phi)}{\partial A}$ in the form $\frac{\partial \ln f(\mathbf{x};A,f,\phi)}{\partial A} = I(A)(g(\mathbf{x})-A) = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}(\frac{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} - A)$. The MVU is thus given by $g(\mathbf{x}) = \frac{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{x}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}$.

(Solution d) The MVU consists of a weighted linear combination of the samples in $\mathbf{x}$ and as these are all Gaussian distributed, the MVU is Gaussian distributed with mean A (it should be unbiased and $E[g(\mathbf{x})] = A\frac{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}} = A$) and variance equal to the inverse of the fisher information: $\hat{A}_{MVU} \sim N(A, \frac{1}{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}})$

(Solution e) If an MVU estimator exists, the MLE procedure will produce. Hence, in this case, as the MVU exists, the MLE will be identical.