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

# ET 4386 Estimation and Detection - Answers

April 12th 2018, 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)

**(2 pts) (a)** $\frac{\partial \log f(x;\theta)}{\partial\theta} = \frac{\partial -\log(\theta)+x\log(1-1/\theta)}{\partial\theta} = -\frac{1}{\theta} + \frac{x}{\theta(\theta-1)}$.

$E\left[\frac{\partial \ln f(x;\theta)}{\partial\theta}\right] = E\left[-\frac{1}{\theta} + \frac{x}{\theta(\theta-1)}\right] = -\frac{1}{\theta} + \frac{1}{\theta} = 0.$

**(1 pts) (b)** We calculate in question a) that $\frac{\partial \log f(x;\theta)}{\partial\theta} = -\frac{1}{\theta} + \frac{x}{\theta(\theta-1)}$. Calcualting $\frac{\partial \log f(x;\theta)}{\partial\theta} = 0$ we get $\frac{1}{\theta} = \frac{x}{\theta(\theta-1)}$ => $(\theta-1) = x$, and thus we obtain $\hat{\theta}_{MLE} = x+1$.

**(3 pts) (c)** We calculate in question a) that $\frac{\partial \log f(x;\theta)}{\partial\theta} = -\frac{1}{\theta} + \frac{x}{\theta(\theta-1)}$. We can then calculate $\frac{\partial^2 \log f(x;\theta)}{\partial\theta^2} = \frac{1}{\theta^2} - \frac{x(2\theta-1)}{\theta^2(\theta-1)^2}$. $I(\theta) = -E\left[\frac{\partial^2 \log f(x;\theta)}{\partial\theta^2}\right] = -E\left[\frac{1}{\theta^2} - \frac{x(2\theta-1)}{\theta^2(\theta-1)^2}\right] = -\frac{1}{\theta^2} + \frac{(2\theta-1)}{\theta^2(\theta-1)} = \frac{1}{\theta(\theta-1)}$.

The CRLB is given by $\text{var}(\hat{\theta}) \geq \frac{1}{I(\theta)} = \theta(\theta-1)$.

**(2 pts) (d)** $\text{var}[\hat{\theta}_{MLE}] = \text{var}[x+1] = \text{var}[x] = \theta(\theta-1)$.

**(2 pts) (e)** The variance of the MLE estimator equals the CRLB. The MLE estimator thus equals the MVU.

Alternatively, we can try to write $\frac{\partial \log f(x;\theta)}{\partial\theta} = -\frac{1}{\theta} + \frac{x}{\theta(\theta-1)} = I(\theta)(\hat{\theta}_{MVU} - \theta)$

$\frac{1}{\theta(\theta-1)}(x-(\theta-1)) = \frac{1}{\theta(\theta-1)}(x+1-\theta))$. From this we can also conclude that $\hat{\theta}_{MVU} = x+1$

## Question 2 (10 points)

**(2,5 pts) (a)** $f(\mathbf{x};A) = \frac{1}{2\pi\sqrt{\det(\mathbf{C})}} \exp\left[-\frac{1}{2}\left(\mathbf{x} - A\mathbf{1}\right)^T \mathbf{C}^{-1}\left(\mathbf{x} - A\mathbf{1}\right)\right]$

$\frac{\partial \log f(\mathbf{x};A)}{\partial A} = \mathbf{x}^T\mathbf{C}^{-1}\mathbf{1} - A\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}$

$\frac{\partial^2 \log f(\mathbf{x};\mathbf{A})}{\partial A^2} = -\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}$

$\text{var}[\hat{A}] \geq \frac{1}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} = \frac{\sigma^2(1-\rho^2)}{2-2\rho} = \frac{\sigma^2}{2}(1+\rho)$

**(2,5 pts) (b)** From $\frac{\partial \log f(\mathbf{x};\mathbf{A})}{\partial A} = \mathbf{x}^T\mathbf{C}^{-1}\mathbf{1} - A\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}$ we can write

$$\frac{\partial f(\mathbf{x};\mathbf{A})}{\partial A} = \mathbf{x}^T\mathbf{C}^{-1}\mathbf{1} - A\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1} = \underbrace{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}}_{I(A)}\left(\underbrace{\left(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}\right)^{-1}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{1}}_{\hat{A}} - A\right)$$

The MVU estimator is thus given by

$$\hat{A} = \left(\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}\right)^{-1}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{1} = \frac{x[0] + x[1]}{2}.$$

**(2,5 pts) (c)** For $\rho = -1$ we obtain $\text{var}\hat{A} \geq 0$. Setting $\rho = -1$, means that $w[0]$ and $w[1]$ have perfect negative correlation. Summing the two samples $x[1]$ and $x[0]$ will perfetly cancel the noise. The MVU estimator that reaches this bound has thus no variance as all the noise is perfectly cancelled.

**(2,5 pts) (d)** With this prior we obtain $\frac{\partial \log f(\mathbf{x};\mathbf{A}) + \log f(A)}{\partial A} = \mathbf{x}^T\mathbf{C}^{-1} - \lambda + A\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}$.

$\hat{A}_{MAP} = \frac{\mathbf{x}^T\mathbf{C}^{-1}\mathbf{1} - \lambda}{\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}} = \frac{\mathbf{x}^T\mathbf{C}^{-1}\mathbf{1} - \lambda}{\frac{2}{\sigma^2(1+\rho)}} = \frac{\frac{x[0]+x[1]}{\sigma^2(1+\rho)} - \lambda}{\frac{2}{\sigma^2(1+\rho)}} = \frac{x[0]+x[1]-\sigma^2(1+\rho)\lambda}{2}$

## Question 3 (10 points)

**(2.5 pts) (a)**

$$T(x) = x \geq \frac{\log\left(\lambda\frac{\theta_1}{\theta_0}\right)}{\log\left(\frac{1-\frac{1}{\theta_1}}{1-\frac{1}{\theta_0}}\right)} = \lambda'$$

**(2.5 pts) (b)** $P_{fa} = P(T(x \geq \lambda' | \mathcal{H}_0)) = \frac{1}{\theta_0}\sum_{x=\lambda'}^{\infty}(1-\frac{1}{\theta_0})^x = (1-\frac{1}{\theta_0})^{\lambda'}$. So, $\lambda' = \log(P_{fa})/\log(1-\frac{1}{\theta_0})$.

**(2.5 pts) (c)** $P_D = P(T(x \geq \lambda' | \mathcal{H}_1)) = \frac{1}{\theta_1}\sum_{x=\lambda'}^{\infty}(1-\frac{1}{\theta_1})^x = (1-\frac{1}{\theta_1})^{\lambda'} = (1-\frac{1}{\theta_1})^{\log(P_{fa})/\log(1-\frac{1}{\theta_0})}$.

**(2.5pts) (d)** We now replace $\theta_1$ with its MLE. $\hat{\theta}_{1,MLE} = x+1$. The GRLT is then given by

$$T(x) = \log(\frac{1}{x+1}) + x\log(\frac{x}{x+1}) - x\log(1-\frac{1}{\theta_0}) \geq \log\frac{\lambda}{\theta_0}$$

## Question 4 (10 points)

**(2 pts) (a)** $T(\mathbf{x}) = \sum_{n=0}^{N-1} x[n]s[n] > \gamma'$

$E(T;\mathcal{H}_0) = E\left[\sum_{n=0}^{N-1} w[n]s[n]\right] = 0$

$E(T;\mathcal{H}_1) = E\left[\sum_{n=0}^{N-1}(s[n]+w[n])s[n]\right] = \alpha$

$\text{Var}(T;\mathcal{H}_0) = \text{Var}\left[\sum_{n=0}^{N-1} w[n]s[n]\right] = \sum_{n=0}^{N-1} \text{Var}\left[w[n]\right] s^2[n] = \sigma^2 \sum_{n=0}^{N-1} s^2[n] = \sigma^2\alpha$

$\text{Var}(T;\mathcal{H}_1) = \text{Var}\left[\sum_{n=0}^{N-1}(s[n]+w[n])s[n]\right] = \sum_{n=0}^{N-1} \text{Var}\left[(s[n]+w[n])\right] s^2[n] = \sigma^2 \sum_{n=0}^{N-1} s^2[n] = \sigma^2\alpha$

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0,\sigma^2\alpha) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\alpha,\sigma^2\alpha) & \text{under } \mathcal{H}_1. \end{cases}$$

**(2 pts) (b)** $P_{fa} = P(T(x \geq \lambda'|\mathcal{H}_0)) = Q\left(\frac{\gamma'}{\sqrt{\sigma^2\alpha}}\right)$. So $\gamma' = \sqrt{\sigma^2\alpha}Q^{-1}\left(P_{fa}\right)$.

$P_D = P(T(x \geq \lambda'|\mathcal{H}_1)) = Q\left(\frac{\gamma'-\alpha}{\sqrt{\sigma^2\alpha}}\right) = Q\left(\frac{\sqrt{\sigma^2\alpha}Q^{-1}(P_{fa})-\alpha}{\sqrt{\sigma^2\alpha}}\right) = Q\left(Q^{-1}\left(P_{fa}\right) - \frac{\sqrt{\alpha}}{\sqrt{\sigma^2}}\right)$

**(2 pts) (c)** The detection performance depends on the SNR $\alpha/\sigma^2$, and thus on the total energy $\alpha$. The shape of $s[n]$ thus does not matter. As long as the energy is kept the same, all signals will have the same detection performance.

**(1 pts) (d)** $T(\mathbf{x}) = \mathbf{x}^T\mathbf{C}^{-1}\mathbf{s} > \gamma'$

$E(T;\mathcal{H}_0) = E\left[\mathbf{w}^T\mathbf{C}^{-1}\mathbf{s}\right] = 0$

$E(T;\mathcal{H}_1) = E\left[(\mathbf{s}+\mathbf{w})^T\mathbf{C}^{-1}\mathbf{s}\right] = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$

$\text{Var}(T;\mathcal{H}_0) = E\left[\left(\mathbf{w}^T\mathbf{C}^{-1}\mathbf{s}\right)^2\right] = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$

$\text{Var}(T;\mathcal{H}_1) = E\left[\left(\mathbf{x}^T\mathbf{C}^{-1}\mathbf{s} - E\left[\mathbf{x}^T\mathbf{C}^{-1}\mathbf{s}\right]\right)^2\right] = E\left[\left(\mathbf{w}^T\mathbf{C}^{-1}\mathbf{s}\right)^2\right] = \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0,\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s},\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}) & \text{under } \mathcal{H}_1. \end{cases}$$

**(3 pts) (e)** The detection probability is now given by $P_D = Q\left(Q^{-1}\left(P_{fa}\right) - \sqrt{\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}}\right)$. Increasing $\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$ will thus lead to an increased detection performance. What is the optimal $s[n]$ for $\sum_{n=0}^{N-1} s^2[n] = \alpha$? Notice that in this case, $\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s} = \sum_{n=0}^{N-1} \frac{s^2[n]}{\sigma_n^2}$

1.

$$\max_{\mathbf{s}} \quad \mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$$
$$s.t. \quad \mathbf{s}^T\mathbf{s} = \alpha.$$

The Lagrangian is given by $L = \sum_{n=0}^{N-1} \frac{s^2[n]}{\sigma_n^2} - \lambda \left( \sum_{n=0}^{N-1} s^2[n] - \alpha \right)$ Taking partial derivatives we get

$$\frac{\partial L}{\partial s[k]} = \frac{2s[k]}{\sigma_k^2} - 2\lambda s[k] = 2s[k](\frac{1}{\sigma_k^2} - \lambda) = 0.$$

$\lambda$ is constant for all samples. So for one $k$, say $k = j$, we have $\lambda = \frac{1}{\sigma_j^2}$. For all other $k \neq j$, $s[k] = 0$. The signal that maximizes the detection probability is therefore given by a signal that is zero always except for one sample $j$. At sample $j$, the signal gets energy $\alpha$. This should be that $j$ where $\sigma_n^2$ has the minimum value, in order to maximize $\mathbf{s}^T\mathbf{C}^{-1}\mathbf{s}$