E1 244: Detection and Estimation Theory
Spring 2011 – Test 2 Solutions

1. Let the observations $Y_1, Y_2, \ldots, Y_N$ be given by

$$Y_n = As_n + W_n$$

where s_n is a known sequence, and W_n is i.i.d. $\mathcal{N}(0, \sigma^2)$.

(a) Assuming σ^2 is known:

i. Find the ML estimator of A.

Solution:

Let $\mathbf{y} = (Y_1, \ldots, Y_N)$, $\mathbf{s} = (S_1, \ldots, S_N)$ and $\mathbf{w} = (W_1, \ldots, W_N)$. Note that $\mathbf{w} \sim \mathcal{N}(0, \sigma^2 I_{N\times N})$. Now, the conditional distribution of $\mathbf{y}$ is given by:

$$f_{\mathbf{y}|A,\mathbf{s}}(\mathbf{y}|A, \mathbf{s}) = \frac{1}{(\sqrt{2\pi\sigma^2})^N} \exp\left\{-\frac{(\mathbf{y} - A\mathbf{s})^T(\mathbf{y} - A\mathbf{s})}{2\sigma^2}\right\} \tag{1}$$

Taking the derivative of the log $f_{\mathbf{y}|A,\mathbf{s}}(\mathbf{y}|A, \mathbf{s})$ and equating to zero gives:

$$2(\mathbf{s}^T\mathbf{y}) - 2A(\mathbf{s}^T\mathbf{s}) = 0, \tag{2}$$

which implies that

$$\hat{A} = \frac{\mathbf{s}^T\mathbf{y}}{\mathbf{s}^T\mathbf{s}}.$$

ii. Find the PDF of the ML estimator of A.

Solution:

Clearly, a linear combination of Gaussian random variables is a Gaussian random variable. Therefore, it suffices to find the mean and the variance of $\hat{A}$. It is easy to show that $\mathbb{E}\hat{A} = A$, and $\text{var}(\mathbf{s}^T\mathbf{y}) = \sigma^2(\mathbf{s}^T\mathbf{s})$. This implies that $\text{var}\hat{A} = \frac{\sigma^2}{\mathbf{s}^T\mathbf{s}}$.

iii. Does the asymptotic normality theorem hold?

Solution:

Since the estimate is Gaussian, $\sqrt{N}(\hat{A} - A) \sim \mathcal{N}(0, \frac{N\sigma^2}{\mathbf{s}^T\mathbf{s}})$. In fact, this is Gaussian for every N. Now, for $N = 1$, we will find the fisher information i_A, and show that $i_A = \frac{\mathbf{s}^T\mathbf{s}}{\sigma^2}$. Since $i_A = -\mathbb{E}\frac{\partial^2 f_{\mathbf{y}|A,\mathbf{s}}(\mathbf{y}|A,\mathbf{s})}{\partial A^2}$, differentiating (2) again with respect to A, we get $\frac{\sigma^2}{\mathbf{a}^T\mathbf{s}}$, which is a constant. Therefore, $i_A = \frac{\sigma^2}{\mathbf{a}^T\mathbf{s}}$. Clearly, the variance is the reciprocal of the Fisher information i_A. This proves the asymptotic normality of $\hat{A}$.

(b) Assuming A is known:

i. Find the ML estimator of σ^2 (or should we ask for σ?).
Solution:
Similar to part(a)-(i), it can be shown that the ML estimate of σ^2 given A is

$$\hat{\sigma^2} = \frac{1}{N}(\mathbf{y} - A\mathbf{s})^T(\mathbf{y} - A\mathbf{s}).$$

ii. Find the PDF of the ML estimator of σ.
Solution:
Clearly, $(\mathbf{y} - A\mathbf{s})^T(\mathbf{y} - A\mathbf{s}) = \sum_{i=1}^N (Y_i - AS_i)^2$. Since $(Y_i - AS_i) \sim \mathcal{N}(0, \sigma^2)$, $\hat{\sigma^2}$ is the sum of square of Gaussian random variables with zero mean and a variance of $\frac{\sigma^2}{N}$, which is chi-square distributed with $N/2$ degrees of freedom. Specifically, we have the following:

$$f_{\hat{\sigma^2}|A,\mathbf{s}}(x) = \left(\frac{N}{\sqrt{2\pi\sigma^2}}\right)^{N/2} \frac{1}{\Gamma\left(\frac{N}{2}\right)} x^{N/2-1} e^{-\frac{Nx}{2\sigma^2}}.$$

iii. Does the asymptotic normality theorem hold?
Solution:
Yes. The Fisher information per sample is given by,

$$i_{\sigma^2} = -\mathbb{E}\left(\frac{\partial \log f_{Y_i|A,S_i}(Y_i|A,S_i)}{\partial\sigma^2}\right)^2 = \frac{1}{2\sigma^4},$$

where in the above, we have used $\mathbb{E}W_i^4 = 3\sigma^4$ to obtain the result. Now, by central limit theorem, we have:

$$\frac{\frac{1}{N}(\mathbf{y} - A\mathbf{s})^T(\mathbf{y} - A\mathbf{s}) - \sigma^2}{\sqrt{\frac{2\sigma^4}{N^2}}} \to \mathcal{N}(0, 1), \quad (3)$$

$$\Rightarrow \sqrt{N}\left(\frac{1}{N}(\mathbf{y} - A\mathbf{s})^T(\mathbf{y} - A\mathbf{s}) - \sigma^2\right) \to \mathcal{N}(0, \frac{2\sigma^4}{N}). \quad (4)$$

Thus, the result follows.

2. Let $Y_i \sim f(y|\theta)$, $i = 1, \cdots, n$. For the following models, derive the ML estimates of θ, the CRB, and verify if the ML estimators are efficient.

(a) Poisson distribution with mean θ, i.e.,

$$f(\mathbf{y}|\theta) = \prod_{i=1}^n \frac{e^{-\theta}\theta^{y_i}}{(y_i!)}$$

Solution: The above pdf results in the log-likelihood function

$$\ln f(\mathbf{y}|\theta) = -n\theta + (\ln\theta)\sum_{i=1}^n y_i - \sum_{i=1}^n \ln(y_i!).$$

Differentiating the above, it is easy to show that

$$\hat{\theta}_{ML} = \arg\max_{\theta} \ln f(\mathbf{y}|\theta) = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

The Fisher information and CRB are found by

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} f(\mathbf{y}|\theta)\right] = \frac{1}{\theta^2}\sum_{i=1}^{n} E\left[y_i\right] = \frac{n}{\theta},$$

and

$$\text{CRB} = \frac{\theta}{n}$$

The variance of the ML estimator can be evaluated as

$$\text{Var}_\theta\left[\hat{\theta}_{ML}\right] = E[(\hat{\theta}_{ML}^2 - \theta)^2] = \frac{1}{n^2}((n^2-n)\theta^2 + n(\theta+\theta^2)) - 2\theta^2 + \theta^2 = \frac{\theta}{n}.$$

Thus the ML estimator achieves the CRB.

(b) Exponential distribution with mean $\sqrt{\theta}$, i.e.,

$$f(\mathbf{y}|\theta) = \prod_{i=1}^{n} \frac{e^{-y_i/\sqrt{\theta}}}{\sqrt{\theta}}.$$

Hint for part (c): $\sum_{i=1}^{n} y_i$ has an Erlang distribution with shape parameter n and scale parameter $1/\sqrt{\theta}$, hence

$$E_\theta\left\{\left[\sum_{i=1}^{n} y_i\right]^2\right\} = \theta n(n+1)$$

and

$$E_\theta\left\{\left[\sum_{i=1}^{n} y_i\right]^4\right\} = \theta^2 n(n+1)(n+2)(n+3)$$

Solution: The log-likelihood function is given by

$$\ln f(\mathbf{y}|\theta) = \sum_{i=1}^{n}\left(\frac{1}{\sqrt{\theta}}e^{-\frac{y_i}{\sqrt{\theta}}}\right) = -\frac{n}{2}\log\theta - \frac{1}{\sqrt{\theta}}\sum_{i=1}^{n} y_i.$$

The ML estimator is found by differentiating w.r.t. θ as

$$\hat{\theta}_{ML} = \arg\max_{\theta} \ln f(\mathbf{y}|\theta) = \left[\frac{\sum_{i=1}^{n} y_i}{n}\right]^2.$$

The Fisher Information and the CRB are found by

$$I(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2} f(\mathbf{y}|\theta)\right] = \frac{n}{4\theta^2}, \quad \text{CRB} = \frac{4\theta^2}{n}.$$

The MSE of the estimator can be found using the given hints as

$$E\left[(\hat{\theta}_{ML}-\theta)^2\right] = E\left[\hat{\theta}_{ML}^2\right] - 2\theta E\left[\hat{\theta}_{ML}\right] + \theta^2 = \theta^2\left(\frac{4}{n}+\frac{11}{n^2}+\frac{6}{n^3}\right).$$

The estimator is thus not efficient but it is asymptotically efficient (since the higher order terms decay faster than n.

3. In this problem, the equations for extended Kalman (EKF) filter when state equation and/or observation equation are nonlinear, are derived.

$$\begin{aligned} \mathbf{x}_n &= f(\mathbf{x}_{n-1}) + G\mathbf{u}_n \\ \mathbf{y}_n &= h(\mathbf{x}_n) + \mathbf{v}_n \end{aligned}$$

$f()$ and $h()$ are functions of appropriate dimensions.

(a) First consider the modified state model that has known deterministic input,

$$\mathbf{x}_n = F_{n-1}\mathbf{x}_{n-1} + G\mathbf{u}_n + \mathbf{t}_n$$

where $\mathbf{t}_n$ is known. The observation equation is

$$\mathbf{y}_n = H_n\mathbf{x} + \mathbf{v}_n$$

Determine the equations for this modified state model.
Hint: Write $\mathbf{x}_n = \mathbf{x}'_n + \mathbb{E}[\mathbf{x}_n]$, where $\mathbf{x}'_n$ is the value of $\mathbf{x}_n$ when $\mathbf{t}_n = 0$. Then find the Kalman filter for zero mean signal filter $\mathbf{x}'_n$.
Also, let $\mathbf{y}'_n = \mathbf{y}_n - H_n\mathbb{E}[\mathbf{x}_n]$, to get observation equation in usual form.
Solution:
We can write the equations as,

$$\begin{aligned} \mathbf{x}'_n &= F_{n-1}\mathbf{x}'_{n-1} + G\mathbf{u}_n \\ \mathbf{y}'_n &= H_n\mathbf{x}' + \mathbf{v}_n \end{aligned}$$

Also the mean satisfies

$$\mathbb{E}[\mathbf{x}_n] = F_{n-1}\mathbb{E}[\mathbf{x}_{n-1}] + \mathbf{t}_n \tag{5}$$

Then the MMSE estimate of $\mathbf{x}_n$ is,

$$\begin{aligned} \hat{\mathbf{x}'}_{n|n-1} &= \hat{\mathbf{x}}_{n|n-1} - \mathbb{E}[\mathbf{x}_n] \\ \hat{\mathbf{x}'}_{n-1|n-1} &= \hat{\mathbf{x}}_{n-1|n-1} - \mathbb{E}[\mathbf{x}_{n-1}] \end{aligned}$$

Also we have prediction equation as,

$$\hat{\mathbf{x}'}_{n|n-1} = F_n\hat{\mathbf{x}'}_{n-1|n-1}$$

or

$$\hat{\mathbf{x}}_{n|n-1} - \mathbb{E}[\mathbf{x}_n] = F_n(\hat{\mathbf{x}}_{n-1|n-1} - \mathbb{E}[\mathbf{x}_{n-1}])$$

From (5) this reduces to

$$\hat{\mathbf{x}}_{n|n-1} = F_n \hat{\mathbf{x}}_{n-1|n-1} + \mathbf{t}_n$$

For the correction equation,

$$\hat{\mathbf{x}'}_{n|n} = \hat{\mathbf{x}'}_{n|n-1} + K_n(\mathbf{y}'_n - H_n \hat{\mathbf{x}'}_{n|n-1})$$

Which reduces to usual expression,

$$\hat{\mathbf{x}}_{n|n} = \hat{\mathbf{x}}_{n|n-1} + K_n(\mathbf{y}_n - H_n \hat{\mathbf{x}}_{n|n-1})$$

(b) Now, first linearize using a first order Taylor expansion about the estimate of $\mathbf{x}_{n-1}$,

$$\begin{aligned} f(\mathbf{x}_{n-1}) &\approx f(\hat{\mathbf{x}}_{n-1|n-1}) \\ &+ \left.\frac{\partial f}{\partial \mathbf{x}_{n-1}}\right|_{\mathbf{x}_{n-1}=\hat{\mathbf{x}}_{n-1|n-1}} (\mathbf{x}_{n-1} - \hat{\mathbf{x}}_{n-1|n-1}) \\ h(\mathbf{x}_n) &\approx h(\hat{\mathbf{x}}_{n|n-1}) \\ &+ \left.\frac{\partial h}{\partial \mathbf{x}_n}\right|_{\mathbf{x}_n=\hat{\mathbf{x}}_{n|n-1}} (\mathbf{x}_n - \hat{\mathbf{x}}_{n|n-1}) \end{aligned}$$

Denote the Jacobians by,

$$\begin{aligned} F_{n-1} &= \left.\frac{\partial f}{\partial \mathbf{x}_{n-1}}\right|_{\mathbf{x}_{n-1}=\hat{\mathbf{x}}_{n-1|n-1}} \\ H_n &= \left.\frac{\partial h}{\partial \mathbf{x}_n}\right|_{\mathbf{x}_n=\hat{\mathbf{x}}_{n|n-1}} \end{aligned}$$

And write down the state and observation equations in form (a).
Solution:
The expressions are,

$$\begin{aligned} \mathbf{x}_n &= F_{n-1}\mathbf{x}_{n-1} + G\mathbf{u}_n + (f(\hat{\mathbf{x}}_{n-1|n-1}) - F_{n-1}\hat{\mathbf{x}}_{n-1|n-1}) \\ \mathbf{y}_n &= H_n\mathbf{x} + \mathbf{v}_n + (h(\hat{\mathbf{x}}_{n|n-1}) - H_n\hat{\mathbf{x}}_{n|n-1}). \end{aligned}$$

We have the modified observation equation

$$\mathbf{y}_n = H_n\mathbf{x} + \mathbf{v}_n + \mathbf{z}_n,$$

with $\mathbf{z}_n$ known. On writing $\mathbf{y}'_n = \mathbf{y}_n - \mathbf{z}_n$, we have expression in usual form.

(c) Combine above two results to get the EKF.
Solution:
The solutions are,

$$\begin{aligned} \hat{\mathbf{x}}_{n|n-1} &= F_n\hat{\mathbf{x}}_{n-1|n-1} + \mathbf{t}_n \\ \hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + K_n(\mathbf{y}_n - \mathbf{z}_n + H_n\hat{\mathbf{x}}_{n|n-1}). \end{aligned}$$

But

$$\begin{aligned}
\mathbf{t}_n &= f(\hat{\mathbf{x}}_{n-1|n-1}) - F_{n-1}\hat{\mathbf{x}}_{n-1|n-1} \\
\mathbf{z}_n &= h(\hat{\mathbf{x}}_{n|n-1}) - H_n\hat{\mathbf{x}}_{n|n-1},
\end{aligned}$$

so that finally

$$\begin{aligned}
\hat{\mathbf{x}}_{n|n-1} &= f(\hat{\mathbf{x}}_{n-1|n-1}) \\
\hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + K_n(\mathbf{y}_n - h(\hat{\mathbf{x}}_{n|n-1})).
\end{aligned}$$