Tutorial 12

1. Consider a measurement vector $\boldsymbol{r} = [r_1 \; r_2 \; r_3]^T$, and its elements are expressed as:

$$r_1 = A + w_1, \; r_2 = A + w_2, \; r_3 = A + w_3$$

where A is the constant to be estimated, while $w_1 \sim \mathcal{N}(0, \sigma_1^2)$, $w_2 \sim \mathcal{N}(0, \sigma_2^2)$ and $w_3 \sim \mathcal{N}(0, \sigma_3^2)$ are noise components and they are independent of each other.

(a) Write down the probability density function (PDF) of r_1.
(b) Compute the covariance matrix of $\boldsymbol{r}$.
(c) Write down the joint PDF of $\boldsymbol{r}$.
(d) Find the maximum likelihood (ML) estimate of A, $\hat{A}$.
(e) Determine the mean and variance of $\hat{A}$.
(f) Suppose $\boldsymbol{r} = [5.1 \;\; 6.2 \;\; 7.3]^T$ while $\sigma_1^2 = 0.1$, $\sigma_2^2 = 1$ and $\sigma_3^2 = 5$. Compute $\hat{A}$ and the variance of $\hat{A}$.

2. Given N measurements of r_n:

$$r_n = \alpha \sin(\omega n + \phi) + w_n, \quad n = 1, \cdots, N$$

where α is the unknown constant to be estimated, ω and ϕ are known constants, and $w_n \sim \mathcal{N}(0, \sigma^2)$, $n = 1, \cdots, N$, are independent.

(a) Write down the joint PDF of $\boldsymbol{r} = [r_1 \ \cdots \ r_N]^T$.
(b) Find the ML estimate of α, $\hat{\alpha}$.
(c) Determine the mean and variance of $\hat{\alpha}$.

3. Given N measurements of r_n:

$$r_n = \cos(\omega n) + w_n, \quad n = 1, \cdots, N$$

where $\omega \in (0, \pi)$ is the unknown constant frequency of a sinusoid to be estimated, and $w_n \sim \mathcal{N}(0, \sigma^2)$, $n = 1, \cdots, N$, are independent.

(a) Is it a linear or nonlinear model?
(b) Write down the joint PDF of $\boldsymbol{r} = [r_1 \ \cdots \ r_N]^T$.
(c) Express the ML estimate of ω, $\hat{\omega}$, with the use of a cost function.

Solution

1(a)

The PDF of r_1 is:

$$p(r_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(r_1 - A)^2}$$

1(b)

The covariance matrix is:

$$\boldsymbol{C} = \mathbb{E}\{(\boldsymbol{r} - A\mathbf{1})(\boldsymbol{r} - A\mathbf{1})^T\} = \mathbb{E}\{\boldsymbol{w}\boldsymbol{w}^T\} = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

where

$$\boldsymbol{w} = [w_1 \ w_2 \ w_3]^T$$

1(c)

The joint PDF of $\boldsymbol{r}$ has the form of:

$$p(\boldsymbol{r}) = \frac{1}{(2\pi)^{N/2}|\boldsymbol{C}|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{r}-A\mathbf{1})^T \boldsymbol{C}^{-1}(\boldsymbol{r}-A\mathbf{1})}$$

Since:

$$|\boldsymbol{C}| = \sigma_1^2\sigma_2^2\sigma_2^2 \qquad \boldsymbol{C}^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} \end{bmatrix}$$

Hence, we have:

$$p(\boldsymbol{r}) = \frac{1}{(2\pi)^{N/2}\sigma_1\sigma_2\sigma_3} e^{-\frac{1}{2}\left(\frac{(r_1-A)^2}{\sigma_1^2}+\frac{(r_2-A)^2}{\sigma_2^2}+\frac{(r_3-A)^2}{\sigma_3^2}\right)}$$

1(d)

Maximizing the likelihood function means minimizing:

$$\frac{(r_1-\tilde{A})^2}{\sigma_1^2}+\frac{(r_2-\tilde{A})^2}{\sigma_2^2}+\frac{(r_3-\tilde{A})^2}{\sigma_3^2} \qquad \text{or} \qquad \left(\boldsymbol{r}-\tilde{A}\mathbf{1}\right)^T \boldsymbol{C}^{-1}\left(\boldsymbol{r}-\tilde{A}\mathbf{1}\right)$$

From (6.10), the solution is:

$$\hat{A}=\frac{\mathbf{1}^T\boldsymbol{C}^{-1}\boldsymbol{r}}{\mathbf{1}^T\boldsymbol{C}^{-1}\mathbf{1}}=\frac{\dfrac{r_1}{\sigma_1^2}+\dfrac{r_2}{\sigma_2^2}+\dfrac{r_3}{\sigma_3^2}}{\dfrac{1}{\sigma_1^2}+\dfrac{1}{\sigma_2^2}+\dfrac{1}{\sigma_3^2}}$$

1(e)

According to (6.11) and (6.12), we have:

$$\mathbb{E}\{\hat{A}\} = A$$

$$\mathrm{var}\left(\hat{A}\right)=(\mathbf{1}^T\boldsymbol{C}^{-1}\mathbf{1})^{-1}=\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}+\frac{1}{\sigma_3^2}\right)^{-1}$$

1(f)
Based on the results in 1(d) and 1(e), we have

$$\hat{A}=\frac{\mathbf{1}^T\boldsymbol{C}^{-1}\boldsymbol{r}}{\mathbf{1}^T\boldsymbol{C}^{-1}\mathbf{1}}=\frac{\dfrac{r_1}{\sigma_1^2}+\dfrac{r_2}{\sigma_2^2}+\dfrac{r_3}{\sigma_3^2}}{\dfrac{1}{\sigma_1^2}+\dfrac{1}{\sigma_2^2}+\dfrac{1}{\sigma_3^2}}=\frac{\dfrac{5.1}{0.1}+\dfrac{6.2}{1}+\dfrac{7.3}{5}}{\dfrac{1}{0.1}+\dfrac{1}{1}+\dfrac{1}{5}}=5.24$$

and

$$\mathrm{var}\left(\hat{A}\right)=\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}+\frac{1}{\sigma_3^2}\right)^{-1}=\left(\frac{1}{0.1}+\frac{1}{1}+\frac{1}{5}\right)^{-1}=0.0893$$

It can be seen that the variance should be less than the variance of each of the r_1, r_2 and r_3.

Note that if $\sigma_2\to\infty$ and $\sigma_3\to\infty$, then $\mathrm{var}(\hat{A})\to\sigma_1^2=0.1$.

2(a)

Since $\{w_n\}$ are IID with variance σ^2, the covariance matrix is:

$$\boldsymbol{C} = \sigma^2 \boldsymbol{I}_N \Rightarrow |\boldsymbol{C}| = \sigma^{2N}, \quad \boldsymbol{C}^{-1} = \sigma^{-2}\boldsymbol{I}_N$$

Let

$$\boldsymbol{a} = [\sin(\omega+\phi) \ \ \sin(2\omega+\phi) \ \ \cdots \ \ \sin(N\omega+\phi)]^T$$

In matrix form, we have:

$$\boldsymbol{r} = \boldsymbol{a}\alpha + \boldsymbol{w}, \quad \boldsymbol{w} = [w_1 \ \ \cdots \ \ w_N]^T$$

The joint PDF of $\boldsymbol{r}$ is then:

$$p(\boldsymbol{r}) = \frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}(\boldsymbol{r}-\alpha\boldsymbol{a})^T(\boldsymbol{r}-\alpha\boldsymbol{a})}$$

2(b)
Applying (6.10), we obtain:

$$\hat{\alpha} = \frac{\boldsymbol{a}^T \boldsymbol{C}^{-1} \boldsymbol{r}}{\boldsymbol{a}^T \boldsymbol{C}^{-1} \boldsymbol{a}} = \frac{\boldsymbol{a}^T \boldsymbol{r}}{\boldsymbol{a}^T \boldsymbol{a}}$$

Alternatively, the solution can be obtained in scalar form by first constructing the least squares cost function:

$$J(\tilde{\alpha}) = (\boldsymbol{r} - \tilde{\alpha}\boldsymbol{a})^T (\boldsymbol{r} - \tilde{\alpha}\boldsymbol{a}) = \sum_{n=1}^{N} [r_n - \tilde{\alpha}\sin(\omega n + \phi)]^2$$

Differentiating it with respect to $\tilde{\alpha}$ and then setting the resultant expression to zero, we have:

$$\left.\frac{\partial J(\tilde{\alpha})}{\partial \tilde{\alpha}}\right|_{\tilde{\alpha}=\hat{\alpha}} = 0$$

$$\Rightarrow 2\sum_{n=1}^{N}[r_n - \hat{\alpha}\sin(\omega n + \phi)] \cdot -\sin(\omega n + \phi) = 0$$

$$\Rightarrow \sum_{n=1}^{N} r_n \sin(\omega n + \phi) = \hat{\alpha}\sum_{n=1}^{N}\sin^2(\omega n + \phi)$$

$$\Rightarrow \hat{\alpha} = \frac{\sum_{n=1}^{N} r_n \sin(\omega n + \phi)}{\sum_{n=1}^{N}\sin^2(\omega n + \phi)}$$

which gives the same result in scalar form.

2(c)

According to (6.11) and (6.12), we have:

$$\mathbb{E}\{\hat{\alpha}\} = \alpha$$

$$\mathrm{var}(\hat{\alpha}) = (\boldsymbol{a}^T \boldsymbol{C}^{-1} \boldsymbol{a})^{-1} = (\sigma^{-2} \boldsymbol{a}^T \boldsymbol{a})^{-1} = \frac{\sigma^2}{\boldsymbol{a}^T \boldsymbol{a}} = \frac{\sigma^2}{\sum_{n=1}^N \sin^2(\omega n + \phi)}$$

Note that substituting $r_n = \alpha \sin(\omega n + \phi) + w_n$ into $\hat{\alpha}$, we get:

$$\hat{\alpha} = \frac{\displaystyle\sum_{n=1}^N [\alpha \sin(\omega n + \phi) + w_n] \sin(\omega n + \phi)}{\displaystyle\sum_{n=1}^N \sin^2(\omega n + \phi)} = \alpha + \frac{\displaystyle\sum_{n=1}^N w_n \sin(\omega n + \phi)}{\displaystyle\sum_{n=1}^N \sin^2(\omega n + \phi)}$$

Taking the expected value of $\hat{\alpha}$, we only need to change w_n to $\mathbb{E}\{w_n\} = 0$, also resulting in:

$$\mathbb{E}\{\hat{\alpha}\} = \alpha$$

3(a)
It is a nonlinear model as the sinusoidal function is nonlinear:

$$\boldsymbol{f}(\omega) = [\cos(\omega) \ \cdots \ \cos(\omega N)]^T$$

which cannot be expressed as a linear form such as $\boldsymbol{a}\omega$.

3(b)
Following the steps in 2(a), the PDF of $\boldsymbol{r}$ is:

$$p(\boldsymbol{r}) = \frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}(\boldsymbol{r}-\boldsymbol{f}(\omega))^T(\boldsymbol{r}-\boldsymbol{f}(\omega))} = \frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}\left(\sum_{n=1}^{N}(r_n-\cos(\omega n))^2\right)}$$

where

$$(\boldsymbol{r} - \boldsymbol{f}(\omega))^T (\boldsymbol{r} - \boldsymbol{f}(\omega)) = \sum_{n=1}^{N} (r_n - \cos(\omega n))^2$$

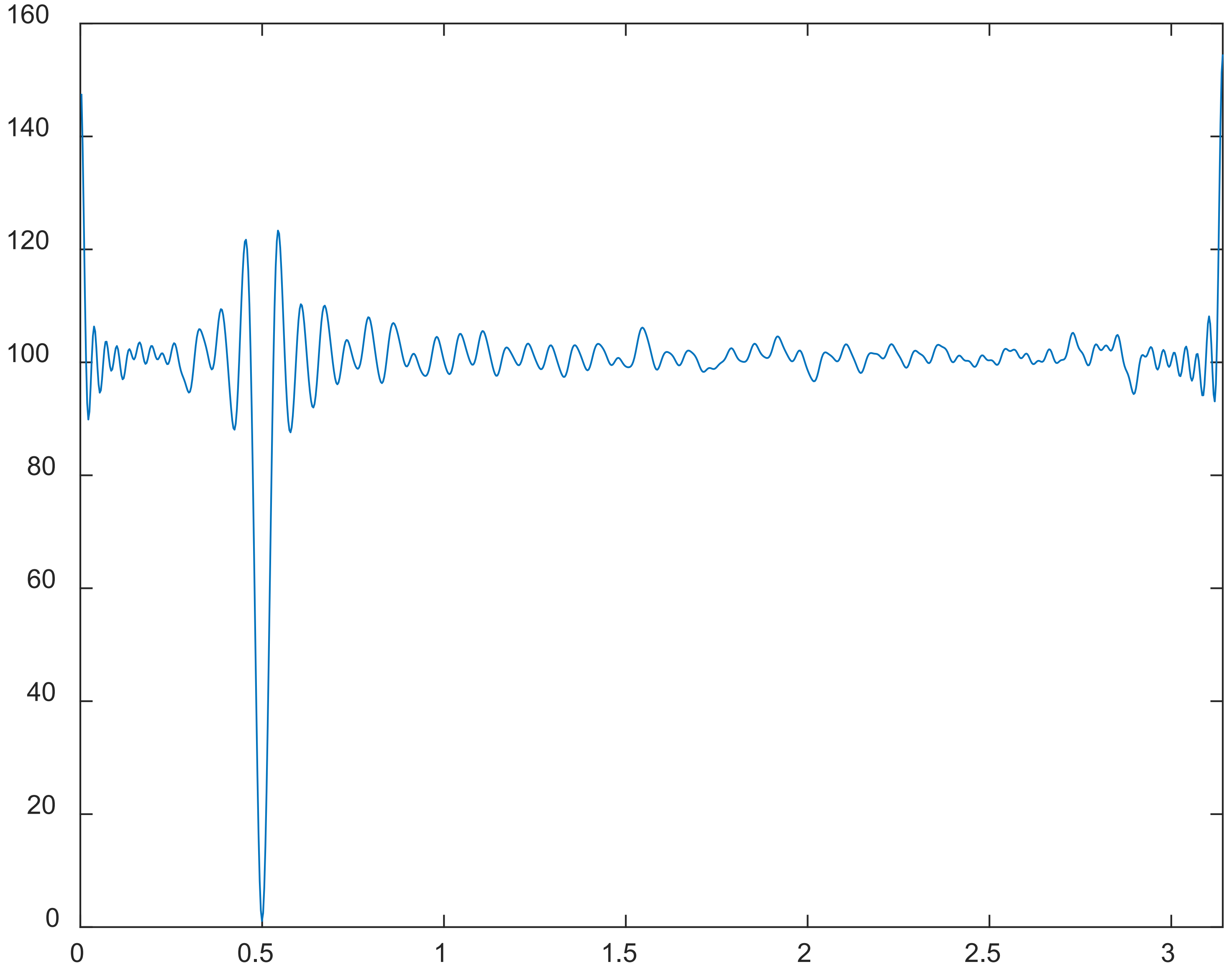
As $\{w_n\}$ are IID, the ML solution is equal to the least squares solution:

$$\hat{\omega} = \arg\min_{\tilde{\omega}} J(\tilde{\omega}), \quad J(\tilde{\omega}) = \sum_{n=1}^{N} (r_n - \cos(\tilde{\omega} n))^2$$

However, it is difficult to find the solution because $J(\tilde{\omega})$ contains multiple minima.

An illustration with $\omega = 0.5$, $N = 100$ and noise power 0.01:

```
w=0.5;
N=100;
n=1:N;
r=cos(w.*n)+0.1*randn(1,N);
for i=1:1000;
f(i)=sum((r-cos(i.*pi./1000.*n)).^2);
end
plot((1:1000)./1000.*pi,f)
axis([0 pi, 0 160])
```


160
140
120
100
80
60
40
20
0
0
0.5
1
1.5
2
2.5
3