6.434J/16.391J

Statistics for Engineers and Scientists — Mar 16

MIT, Spring 2006 — Handout #9

Solution 4

Problem 1: Consider the multi-antenna transmission and reception system. There are $n_T = 2$ transmit and $n_R = 2$ receive antennas. A vector of n_T *iid* data symbols $\mathbf{a} \in \mathcal{A} = \{-1, +1\}^{n_T}$ is transmitted over an AWGN channel. Assume that the symbols are uniformly distributed. The channel coefficients between each transmit and receive antenna pair are grouped in the form of a channel matrix $\mathbf{H}$, which is known in this problem. The observations at the receiver can be written as:

$$\mathbf{r} = \mathbf{H}\mathbf{a} + \mathbf{n}$$

where $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_{n_R})$. $\mathbf{H}$ is specified as the following: $\mathbf{H}_{11} = 0.25, \mathbf{H}_{12} = 0.9, \mathbf{H}_{21} = 0.7, \mathbf{H}_{22} = 0.5$. Implement the above described system in MATLAB for noise variance of $\sigma^2 = 0.25$, 0.5 and 1, respectively. Implement an MMSE estimator to obtain the transmitted vector of symbols. Record the error performance by comparing with the transmitted sequence for each case of σ^2. Be sure to run your system for a sufficient length of samples in each case.

Solution MATLAB code below simulates the system. It implements the L-MMSE estimator for the transmitted symbols. One of the simulation results for the given noise variance values are:

$\sigma_n^2 = 0.25$ mean symbol error = 0.0925

$\sigma_n^2 = 0.50$ mean symbol error = 0.1600

$\sigma_n^2 = 1.00$ mean symbol error = 0.2375

By the way, for a realistic communication system, these errors are quite high. Actually, these values of noise variance are large for such a MIMO system (they yield SNR of 6, 3 and 0 dB respectively). At $\sigma_n^2 = 0.05$ (that is almost 13 dB SNR), the symbol error rate goes below 10^{-3}.

```
---------------------------------------------------------------
% Multi-antenna transmitter and receiver system: MMSE estimator
% Number of transmit and receive antennas:
nT = 2; nR = 2;

% channel coefficient matrix H:
H = [0.25 0.9;0.7 0.5];
```

```
% possible data symbols:
A = [-1;1];

% The noise variance: can change it to any desired value
allVariances = [0.25 0.5 1.0];

RESULT = []; for index = 1:length(allVariances),
    sigma2 = Sigma2(index);
    sampleSize = 1000;
    totalErrors = 0;
    for m=1:sampleSize,
        % Generate transit vectorof  size (nT x 1)
        a = 2*(round(rand(nT,1))-0.5);

        % Generate noise samples
        n = sqrt(sigma2)*randn(nR,1);

        % Receiver sequence is:
        r = H*a + n;

        % L-MMSE of transmitted symbols a, based on observed r:
        a_hat = H.' * inv((H*H.' + sigma2*eye(nR))) * r;

        % Quantize the estimate to be from A:
        a_hatQ = 2*((a_hat > 0)-0.5);

        % Check for errors:
        symError = sum(a_hatQ ~= a);
        totalErrors = totalErrors + symError;
    end
    % For each noise-variance, record sample-mean symbol error
    RESULT = [RESULT; [sigma2 totalErrors/(nT*sampleSize)]];
end
% Display
RESULT

---------------------------------------------------------------
```

Problem 2: Let $X_1, X_2, \ldots, X_n$ be independent random variables, each with density,

$$f(x \,|\, \theta) = \begin{cases} \frac{x}{\theta}\,\mathrm{e}^{-\frac{x^2}{2\theta}} & x > 0; \\ 0, & x \leq 0, \end{cases}$$

where θ is unknown, $\theta \in \Theta \triangleq \{z \,|\, z > 0\}$. (This is called the Rayleigh distribution and has been used to model the fluctuation of the received signal amplitude in wireless transmission.)

(a) Find a maximum likelihood estimator of θ, say, $t_n(X_1, X_2, \ldots, X_n)$.

(b) Use MATLAB to generate $n = 1000$ samples, $X_i = x_i$, of i.i.d. random variables, each with Rayleigh density with parameter $\theta = 2$. Then, plot the estimates $t_m(x_1, x_2, \ldots, x_m)$ as a function of the sample size, m, for $m = 1, 2, \ldots, 1000$. Comment on the figure. Based on the plot, do you think that the estimator is asymptotically unbiased?

Solution

(a) The likelihood and log-likelihood functions are given by

$$\begin{aligned} L(\theta) &= \frac{\prod_{i=1}^n X_i}{\theta^n}\,\mathrm{e}^{-\frac{\sum_{i=1}^n X_i^2}{2\theta}} \\ \ln L(\theta) &= \sum_{i=1}^n \ln X_i - n\ln\theta - \frac{1}{2\theta}\sum_{i=1}^n X_i^2. \end{aligned}$$

We want to find $\theta > 0$ that maximizes the log-likelihood function. The first and second partial derivatives of the log-likelihood function are given by

$$\begin{aligned} \frac{\partial}{\partial\theta}\ln L(\theta) &= -\frac{n}{\theta} + \frac{1}{2\theta^2}\sum_{i=1}^n X_i^2 \\ \frac{\partial^2}{\partial\theta^2}\ln L(\theta) &= \frac{n}{\theta^2} - \frac{1}{\theta^3}\sum_{i=1}^n X_i^2. \end{aligned}$$

Setting the first partial derivative to zero yields a saddle point

$$\theta^* = \frac{\sum_{i=1}^n X_i^2}{2n},$$

which maximizes the log-likelihood function:

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta)\bigg|_{\theta=\theta^*} = -\frac{n}{(\theta^*)^2} < 0.$$

Therefore, the maximum likelihood estimator (MLE) of θ is

$$t_n(X_1, X_2, \dots, X_n) = \frac{\sum_{i=1}^n X_i^2}{2n}.$$

(b) MATLAB command `raylrnd(`$\sqrt{\theta}$`,1,n)` generates n i.i.d. samples of Rayleigh density with a parameter θ. For the data set used to plot Fig. 1, the maximum likelihood estimate approaches $\theta = 2$ when the sample size is large. From the figure, it seems that estimator is asymptotically unbiased.

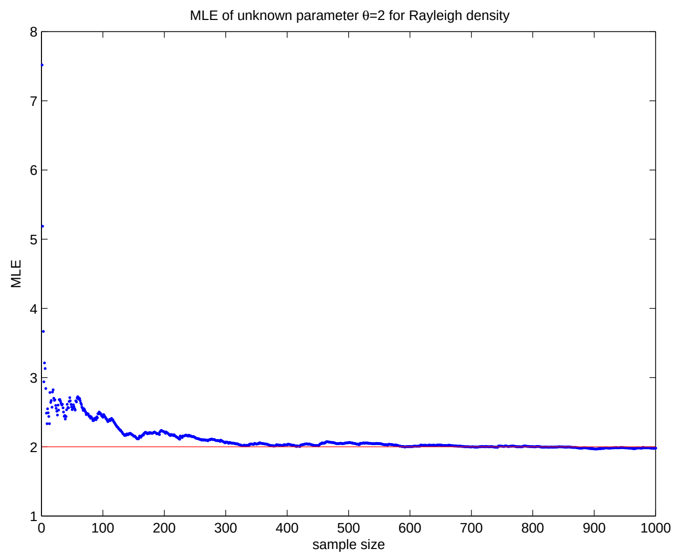

Figure 1: The estimate is close to 2 for a large sample size.

Problem 3: Let $X_1, X_2, \dots, X_n$ be i.i.d. random variables from the Geometric distribution with the probability of success $0 \leq 1/\theta \leq 1$, for an unknown parameter θ:

$$\mathbb{P}\{X_1 = k\} = \left(1 - \frac{1}{\theta}\right)^{k-1} \cdot \frac{1}{\theta}, \quad k = 1, 2, 3, \dots .$$

Let T_n be the maximum likelihood estimator (MLE) of θ based on a random sample of size n.

(a) Fix $\theta = 3$ and use MATLAB to plot a histogram of $\sqrt{n}(T_n - \theta)$, for $n = 500$.

(b) Find the Fisher information, $\mathrm{I}(\theta)$.

Solution First, we derive the MLE of θ. The likelihood and log-likelihood function are

$$\begin{aligned} L(\theta) &= \left[\left(1-\frac{1}{\theta}\right)^{X_1-1}\cdot\frac{1}{\theta}\right]\times\left[\left(1-\frac{1}{\theta}\right)^{X_2-1}\cdot\frac{1}{\theta}\right] \\ &\quad\times\cdots\times\left[\left(1-\frac{1}{\theta}\right)^{X_n-1}\cdot\frac{1}{\theta}\right] \\ &= \left(1-\frac{1}{\theta}\right)^{(\sum_{i=1}^n X_i)-n}\frac{1}{\theta^n} \end{aligned}$$

and

$$\ln L(\theta) = \left[\left(\sum_{i=1}^n X_i\right)-n\right]\ln\left(1-\frac{1}{\theta}\right) - n\ln\theta.$$

The first and second derivatives of the log-likelihood functions with respect to θ are

$$\frac{\partial}{\partial\theta}\ln L(\theta) = \frac{(\sum_{i=1}^n X_i)-n}{\theta^2-\theta} - \frac{n}{\theta} \quad (1)$$

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta) = -\frac{[(\sum_{i=1}^n X_i)-n](2\theta-1)}{\theta^2-\theta} + \frac{n}{\theta^2}. \quad (2)$$

Setting the first derivative to zero and solving for an unknown $\theta \geq 1$ yields a saddle point

$$\theta^* = \frac{\sum_{i=1}^n X_i}{n},$$

which maximizes the log-likelihood function:

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta)\Big|_{\theta=\theta^*} = -\frac{n}{\theta^*(\theta^*-1)} < 0.$$

Therefore, the maximum likelihood estimator of θ is

$$T_n = \frac{\sum_{i=1}^n X_i}{n}.$$

(a) MATLAB code to plot the histogram is shown below:

```
%-----------%
% variables %
%-----------%
m = 10000; % number of points for the histogram
n = 500;  % sample size for each histogram point
```

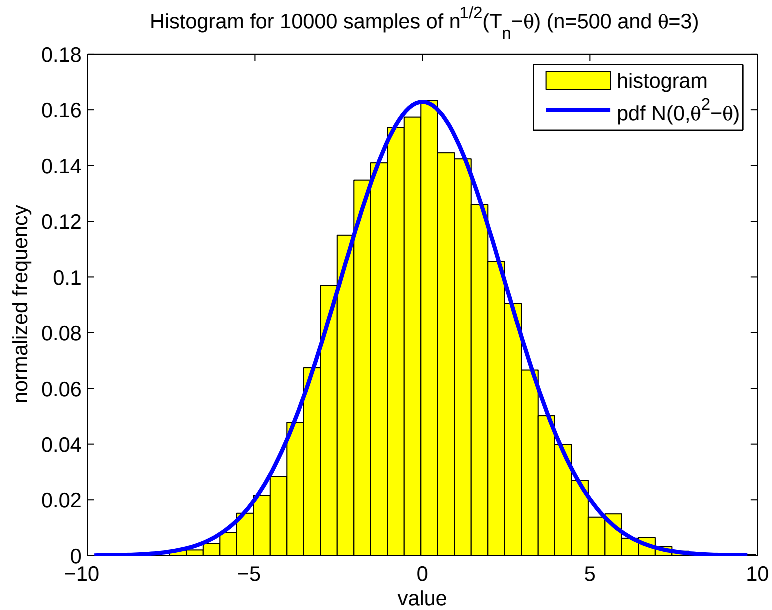

Figure 2: Histogram for problem 3(a) matches the pdf of a normal random variable.

```
theta = 3;

%---------%
% samples %
%---------%
% Note that MATLAB defines geometric distribution to be
%   Pr{ Y = k } = (1-p)^k p , for k=0,1,2,...
% From our homework, we define a r.v. X to be the pmf
%   Pr{ X = l } = (1-p)^(l-1) p, for l=1,2,3,... . ----(*)
% Note that Y+1 has the pmf of (*):
%   Pr{ Y+1 = s } = Pr{ Y=s-1 }
%                 = (1-p)^{s-1} p, for s=1,2,3... .
% Thus, to generate sample of pmf (*), we simply add one to the
% command geornd.
x = 1 + geornd(1/theta, n, m );

%-----------------------------------------%
% derive ML estimates T_n(1), ..., T_n(m) %
```

```
%-----------------------------------------%
% Structure of x:
%   x[1,1]   x[1,2] ... x[1,m]
%   x[2,1]   x[2,2] ... x[2,m]
%   ...
%   x[n,1]   x[n,2] ... x[n,m]
%    |        |           |
%    |        |           V
%    |        |           use this data set for T_n (m)
%    |        V
%    V       use this data set for T_n (2)
%
%  use this data set for T_n (1)

% sum over each column
mle = sum(x)/n;

%-------------------------------------%
% plot histogram and the limiting pdf %
%-------------------------------------%
sigma = sqrt(theta^2-theta); % limiting value of standard dev.

hres = 0.5;                      % resolution of the histogram
t =-4*sigma:hres:4*sigma;        % resolution for histogram
yhist = sqrt(n)*(mle-theta);     % data for histogram plot
[fqcount,xout] = hist(yhist,t);
% plot normalized frequency
% 1 = full width of bars, -y = solid yellow line
bar(xout,fqcount/m/hres, 1); %last argument is width of the bar
                 % edit the figure color to your choice
hold on

val = -4*sigma:0.1:4*sigma; % x-axis, new resolution
y = normpdf( val, 0, sigma );
plot( val, y, '-b', 'Linewidth', 2 ); % so the line is over the histogram

%--------------------------%
% label axes, title, legend %
```

```
%--------------------------%
% add title and lables
legend( 'histogram', 'pdf N(0,\theta^2-\theta)');
xlabel('value');
ylabel('normalized frequency');
txt = ['Histogram for ', num2str(m), ' samples of n^{1/2}...
 (T_n-\theta) (n=', num2str(n),' and \theta=', num2str(theta), ')' ];
title(txt);
```

(b) Let X denote a Geometric random variable with the probability of success $1/\theta$, and let $p(\cdot)$ denote the probability mass function of X:

$$f(k\,|\,\theta) = \left(1 - \frac{1}{\theta}\right)^{k-1} \cdot \frac{1}{\theta}, \qquad k = 1, 2, \ldots .$$

The Fisher's information is given by

$$\begin{aligned}
\mathrm{I}(\theta) &\triangleq \mathbb{E}\left\{\left[\frac{\partial}{\partial\theta} \ln f(X\,|\,\theta)\right]^2\right\} && \text{(the expectation over a random variable } X\text{)} \\
&= \mathbb{E}\left\{\left(\frac{X-1}{\theta^2-\theta} - \frac{1}{\theta}\right)^2\right\} && \text{(from equation (1) with } n = 1\text{)} \\
&= \mathbb{E}\left\{\left(\frac{X-\theta}{\theta^2-\theta}\right)^2\right\} \\
&= \frac{1}{(\theta^2-\theta)^2}\mathbb{E}\left\{(X-\theta)^2\right\} \\
&= \frac{1}{(\theta^2-\theta)^2}\,\mathrm{Var}\,X \\
&= \frac{1}{\theta^2-\theta}.
\end{aligned}$$

Problem 4:

(a) Let $X, X_1, X_2, \ldots, X_n$ be independent random variables, each with the same pdf $f_X(x|\theta)$ for some $\theta \in \Theta$. Define the Fisher information

$$I(\theta) = \mathbb{E}_\theta\left(\frac{\partial}{\partial\theta} \ln f_X(x|\theta)\right)^2$$

Show that

$$\mathbb{E}_\theta\left(\frac{\partial}{\partial\theta} \ln f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n|\theta)\right)^2 = nI(\theta).$$

(b) Let $t_n(X_1, \ldots, X_n)$ be an unbiased estimator of some function of some θ, say $g(\theta)$. Show that

$$Var\left(t_n(X_1, \ldots, X_n)\right) \geq \frac{\left(g'(\theta)\right)^2}{nI(\theta)}$$

Solution

(a) We fist find

$$\begin{aligned}
\mathbb{E}_\theta\left(\frac{\partial}{\partial\theta}\ln f_X(x|\theta)\right) &= \int_{-\infty}^{\infty}\frac{\partial}{\partial\theta}\ln f_X(x|\theta)\cdot f_X(x|\theta)dx \\
&= \int_{-\infty}^{\infty}\frac{\frac{\partial}{\partial\theta}f_X(x|\theta)}{f_X(x|\theta)}f_X(x|\theta)dx \\
&= \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}f_X(x|\theta)dx \\
&= \frac{\partial}{\partial\theta}(1) \\
&= 0
\end{aligned}$$

Next, note that, if random variables V and W are independent, then, for any function of V and W, they are independent, i.e., $\mathbb{E}\left\{g(W)g(V)\right\} = \mathbb{E}\left\{g(W)\right\}\mathbb{E}\left\{g(V)\right\}$. Thus, we have

$$\begin{aligned}
&\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_1,\ldots,X_n}(x_1,\ldots,x_n|\theta)\right)^2 = \mathbb{E}\left(\frac{\partial}{\partial\theta}\sum_{i=1}^{n}\ln f_{X_i}(x_i|\theta)\right)^2 \\
&= \sum_{i=1}^{n}\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}(x_i|\theta)\right)^2 + 2\sum_{i<j}\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}(x_i|\theta)\cdot\frac{\partial}{\partial\theta}\ln f_{X_j}(x_j|\theta)\right) \\
&= \sum_{i=1}^{n}I(\theta) + 0 \\
&= nI(\theta)
\end{aligned}$$

Hence, we conclude $\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_1,\ldots,X_n}(x_1,\ldots,x_n|\theta)\right)^2 = nI(\theta)$.

(b) Since $t_n(X_1, \ldots, X_n)$ is a unbiased estimator of function $g(\theta)$, we have

$$\begin{aligned}
g(\theta) &= \mathbb{E}\left\{t_n(X_1,\ldots,X_n)\right\} \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}t_n(x_1,\ldots,x_n)\left(\prod_{i=1}^{n}f_{X_i}(x_i|\theta)\right)dx_1\ldots dx_n
\end{aligned}$$

From (a), we know that

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(\theta)\left(\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)\prod_{i=1}^{n} f_{X_i}(x_i|\theta)dx_1\ldots dx_n$$
$$= g(\theta)\sum_{i=1}^{n}\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln f_{X_i}(x_i|\theta)\right) \quad = g(\theta)\cdot 0 \quad = 0$$

Then, we have

$$\begin{aligned}
g'(\theta) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} t_n(x_1,\ldots,x_n)\left(\frac{\partial}{\partial\theta}\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)dx_1\ldots dx_n \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} t_n(x_1,\ldots,x_n)\left(\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)\prod_{i=1}^{n} f_{X_i}(x_i|\theta)dx_1\ldots dx_n \\
&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (t_n(x_1,\ldots,x_n)-g(\theta))\left(\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)\prod_{i=1}^{n} f_{X_i}(x_i|\theta)dx_1\ldots dx_n \\
&= \mathbb{E}\left((t_n(X_1,\ldots,X_n)-g(\theta))\left(\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)\right) \\
&\leq \left(\mathbb{E}\left\{t_n(X_1,\ldots,X_n)-\mathbb{E}\left[t_n(X_1,\ldots,X_n)\right]\right\}^2\right)^{\frac{1}{2}}\left(\mathbb{E}\left(\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)^2\right)^{\frac{1}{2}} \\
&= \left(Var(t_n(X_1,\ldots,X_n))\right)^{\frac{1}{2}}\left(nI(\theta)\right)^{\frac{1}{2}}
\end{aligned}$$

Note that the third equality holds since $\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(\theta)\left(\frac{\partial}{\partial\theta}\ln\prod_{i=1}^{n} f_{X_i}(x_i|\theta)\right)$ $\prod_{i=1}^{n} f_{X_i}(x_i|\theta)dx_1\ldots dx_n = 0$, and the inequality is the from the Schwartz inequality. Therefore, we have $Var\left(t_n(X_1,\ldots,X_n)\right) \geq \frac{\left(g'(\theta)\right)^2}{nI(\theta)}$.

Problem 5:

(a) For the following system:

$$\mathbf{r} = \alpha\mathbf{s} + \mathbf{n}$$

where $\mathbf{s}$ is a vector of N symbols (known), $\mathbf{s} \in \{-1,+1\}^N$, $\mathbf{n} = \mathbf{n}_R + j\mathbf{n}_I$ and both $\mathbf{n}_R$ and $\mathbf{n}_I$ are samples of *iid* zero-mean Gaussian noise with variance $\frac{\sigma^2}{2}$. $\alpha \in \mathbb{R}$, is the parameter that we want to estimate.

i. Find the maximum-likelihood estimate of α, that is $\hat{\alpha}_{\text{ML}}$ and show that it is unbiased.

ii. Find the variance of estimation error, $V(\alpha)$.

iii. Find the lower bound on estimation error for unbiased estimators. This is Cramer-Rao Lower Bound (CRLB).

iv. Plot V and CRLB as a function of σ^2 for $\alpha = 2$ and $N = 20$. [Generally, we plot the performance against the so-called signal-to-noise-ratio (SNR). To get the similar plot, make your plots as a function of $\frac{1}{\sigma^2}$].

(b) Repeat the same exercise as part (a) for the unknown θ in the system model: $\mathbf{r} = e^{j\theta}\mathbf{s} + \mathbf{n}$. You can use the function $\phi = \mathcal{A}rg(x)$, which gives the phase of any complex x s.t. $\phi \in [-\pi, \pi]$. You can simulate to find numerical solution, if the closed form solutions are not easily obtained.

Solution

(a) In this part, we can assume $\mathbf{n} = \mathbf{n}_R$ that makes $\mathbf{r}$ also real because the only imaginary part could be due to the noise, $\mathbf{n}_I$ which can be removed before the estimator.

i. We have to find MLE of α. The likelihood function is:

$$L(\alpha) = p(\mathbf{r}|\alpha) = \left(\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2}}}\right)^N e^{-\sum_{k=0}^{N-1}(r_k - \alpha s_k)^2/2\frac{\sigma^2}{2}}$$

So the log-likelihood function is,

$$\ln L(\alpha) = -\frac{N}{2}\ln(\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{k=0}^{N-1}(r_k - \alpha s_k)^2$$

We want to find α that maximizes log-likelihood function. Taking partial derivatives w.r.t. α:

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\ln L(\alpha) &= \frac{2}{\sigma^2}\sum_{k=0}^{N-1}(r_k - \alpha s_k)s_k \\
&= \frac{2}{\sigma^2}\left(\sum_{k=0}^{N-1} r_k s_k - \alpha\sum_{k=0}^{N-1} s_k^2\right) \\
&= \frac{2}{\sigma^2}\left(\sum_{k=0}^{N-1} r_k s_k - \alpha\sum_{k=0}^{N-1} 1\right) \\
&= \frac{2}{\sigma^2}\left(\sum_{k=0}^{N-1} r_k s_k - N\alpha\right) \qquad (3) \\
\frac{\partial^2}{\partial\alpha^2}\ln L(\alpha) &= -\frac{2N}{\sigma^2} < 0
\end{aligned}$$

Since σ^2 is always positive, second derivative is always negative implying that we indeed have the maximum here. Setting equal to zero and solving for α gives:

$$\hat{\alpha}_{\mathrm{MLE}}(\mathbf{r}) = \frac{1}{N}\sum_{k=0}^{N-1} r_k s_k \tag{4}$$

ii.

$$\begin{aligned}
\mathbb{E}\left\{\hat{\alpha}_{\mathrm{MLE}}(\mathbf{r})\right\} &= \mathbb{E}\left\{\frac{1}{N}\sum_{k=0}^{N-1} r_k s_k\right\} \\
&= \frac{1}{N}\sum_{k=0}^{N-1} s_k \mathbb{E}\left\{r_k\right\} \\
&= \frac{1}{N}\sum_{k=0}^{N-1} \alpha s_k^2 \\
&= \frac{\alpha}{N}\sum_{k=0}^{N-1} 1 \\
&= \alpha
\end{aligned} \tag{5}$$

where the second last step is because $s_k \in \{-1, +1\}$. Thus the estimator is unbiased, its mean estimation error is. So, the variance of estimation error is:

$$\begin{aligned}
\mathbb{V}ar\{\hat{\alpha}_{\mathrm{MLE}} - \alpha\} &= \mathbb{E}\left\{(\hat{\alpha}_{\mathrm{MLE}}(\mathbf{r}) - \alpha)^2\right\} \\
&= \mathbb{E}\left\{\left(\frac{1}{N}\sum_{k=0}^{N-1} r_k s_k - \alpha\right)^2\right\} \qquad (6) \\
&= \mathbb{E}\left\{\left(\frac{1}{N}\sum_{k=0}^{N-1} (\alpha s_k^2 + n_k s_k) - \alpha\right)^2\right\} \\
&= \mathbb{E}\left\{\left(\alpha + \frac{1}{N}\sum_{k=0}^{N-1} n_k s_k - \alpha\right)^2\right\} \\
&= \frac{1}{N^2}\mathbb{E}\left\{\left(\sum_{k=0}^{N-1} n_k s_k\right)^2\right\} \\
&= \frac{1}{N^2}\mathbb{E}\left\{\sum_{k=0}^{N-1} n_k^2 s_k^2 + \sum_{i,j:j\neq i}^{N-1} n_i n_j s_i s_j\right\} \\
&= \frac{1}{N^2}\sum_{k=0}^{N-1} \mathbb{E}\left\{n_k^2\right\} + \sum_{i,j:j\neq i}^{N-1} \mathbb{E}\left\{n_i\right\}\mathbb{E}\left\{n_j\right\} s_i s_j \\
&= \frac{\sigma^2}{2N} \qquad (7)
\end{aligned}$$

Note that the estimation error is inversely related with the vector size, which is expected as the same α scales all the components and estimation should improve with the increase in sample size.

iii. To find the lower bound on mean square estimation error (or variance of estimation error, because mean of the error is zero), we find Fisher's information $I_{\mathbf{r}}(\alpha)$, using (3):

$$\begin{aligned} I_{\mathbf{r}}(\alpha) &= \mathbb{E}\left\{\left(\frac{\partial}{\partial\alpha}\ln p(\mathbf{r}|\alpha)\right)^2\right\} \\ &= \frac{4}{\sigma^4}\mathbb{E}\left\{\left(\sum_{k=0}^{N-1} r_k s_k - N\alpha\right)^2\right\} \\ &= \frac{4N^2}{\sigma^4}\mathbb{E}\left\{\left(\frac{1}{N}\sum_{k=0}^{N-1} r_k s_k - \alpha\right)^2\right\} \end{aligned}$$

where we note that expectation is the same as in (6). Therefore, the result of expectation is given by (7).

$$\begin{aligned} I_{\mathbf{r}}(\alpha) &= \frac{4N^2}{\sigma^4}\frac{\sigma^2}{2N} \\ &= \frac{2N}{\sigma^2} \end{aligned}$$

Note that we could have used:

$$\begin{aligned} I_{\mathbf{r}}(\alpha) &= -\mathbb{E}\left\{\frac{\partial^2}{\partial\alpha^2}\ln p(\mathbf{r}|\alpha)\right\} \\ &= \frac{2N}{\sigma^2} \end{aligned}$$

which is the same result. For the numerator part of the CRLB, we already found that this is an unbiased estimator of α. Therefore,

$$\psi(\alpha) = \mathbb{E}\left\{\hat{\alpha}_{\mathrm{MLE}}(\mathbf{r})\right\} = \alpha$$

Thus,

$$\psi'(\alpha) = 1$$

Hence, the lower bound of mean square estimation error is given by:

$$\frac{(\psi'(\alpha))^2}{I_{\mathbf{r}}(\alpha)} = \frac{\sigma^2}{2N} \tag{8}$$

From part (ii), eq. (7), we see that MLE achieves the lower bound. Note that the noise variance for this derivation is $\frac{\sigma^2}{2}$, that brings the constant

2 in the result.

iv. For $N = 20$, the plot of CRLB vs $\frac{1}{\sigma^2}$is given below. Variance of estimation errir by simulating the estimator with sample size 100 is also shown. Figure below shows the mean of the estimate at different variance.

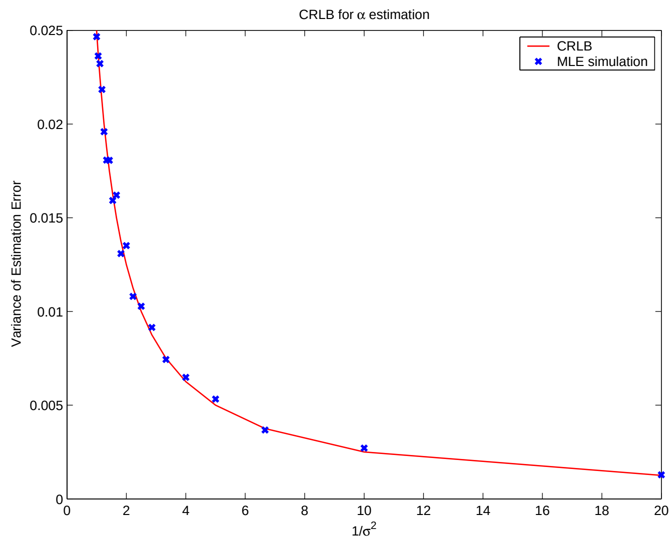

Figure 3: The plot of CRLB and simulated estimator's error variance (sample size 100) as a function of noise variance $\frac{1}{\sigma^2}$.

(b) $\mathbf{r} = e^{j\theta}\mathbf{s} + \mathbf{n}$

i. For the MLE, we first write the likelihood function and the log-likelihood function:

$$\begin{aligned} L(\theta) = p(\mathbf{r}|\theta) &= \left(\frac{1}{\sqrt{\pi\sigma^2}}\right)^N e^{-\frac{1}{\sigma^2}||\mathbf{r}-\mathbf{s}e^{j\theta}||^2} \\ &= \left(\frac{1}{\sqrt{\pi\sigma^2}}\right)^N e^{-\sum_{k=0}^{N-1}|r_k - s_k e^{j\theta}|^2/\sigma^2} \\ \ln L(\theta) &= -\frac{N}{2}\ln(\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{k=0}^{N-1}|r_k - s_k e^{j\theta}|^2 \end{aligned}$$

Note that r_k is complex. In the equations below, $\Re$ denotes real part, $\Im$

denotes imaginary part and * indicates complex conjugate.

$$\begin{aligned}
\ln L(\theta) &= -\frac{N}{2}\ln(\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{k=0}^{N-1}(|r_k|^2 + s_k^2 - 2\Re\{r_k^* s_k e^{j\theta}\}) \\
&= -\frac{N}{2}\ln(\pi\sigma^2) - \frac{1}{\sigma^2}(\sum_{k=0}^{N-1}|r_k|^2 + \sum_{k=0}^{N-1}s_k^2 - \sum_{k=0}^{N-1}2\Re\{r_k^* s_k e^{j\theta}\}) \\
&= -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{1}{\sigma^2}\sum_{k=0}^{N-1}|r_k|^2 - \frac{N}{\sigma^2} + \frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Re\{r_k^* s_k e^{j\theta}\}
\end{aligned}$$

Taking partial derivatives w.r.t. θ:

$$\frac{\partial}{\partial\theta}\ln L(\theta) = -\frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Im\{r_k^* s_k e^{j\theta}\} \tag{9}$$

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta) = -\frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Re\{r_k^* s_k e^{j\theta}\} \tag{10}$$

Equating (9) to 0 and solving for θ,

$$\begin{aligned}
\frac{\partial}{\partial\theta}\ln L(\theta) &= -\frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Im\{r_k^* s_k e^{j\theta}\} \\
&= -\frac{2}{\sigma^2}\Im\left\{e^{j\theta}\sum_{k=0}^{N-1}r_k^* s_k\right\} \\
&= 0
\end{aligned}$$

Imaginary part 0 demands the angle of the term in summation to be $-\theta$. Therefore, the required MLE, $\hat{\theta}_{\mathrm{MLE}(\mathbf{r})}$ is:

$$\begin{aligned}
\hat{\theta}_{\mathrm{MLE}}(\mathbf{r}) &= -\mathcal{A}rg\left(\sum_{k=0}^{N-1}r_k^* s_k\right) \\
&= -\mathcal{A}rg\left(\left(\sum_{k=0}^{N-1}r_k s_k\right)^*\right) \\
&= \mathcal{A}rg\left(\sum_{k=0}^{N-1}r_k s_k\right) \qquad (11) \\
&= \mathcal{A}rg\left(\mathbf{r}^T\mathbf{s}\right) \qquad (12)
\end{aligned}$$

which indeed maximizes the log-likelihood function because plugging $\hat{\theta}_{\mathrm{MLE}}$ in (10), we get:

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta^2}\ln L(\hat{\theta}) &= -\frac{2}{\sigma^2}\Re\{e^{j\hat{\theta}}\sum_{k=0}^{N-1} r_k^* s_k\} \\
&= -\frac{2}{\sigma^2}\Re\left\{\left|\sum_{k=0}^{N-1} r_k^* s_k\right|\right\} < 0
\end{aligned}
$$

ii.

$$
\begin{aligned}
\mathbb{E}\left\{\hat{\theta}\right\} &= \mathbb{E}\left\{\mathcal{A}rg\left(\sum_{k=0}^{N-1} r_k s_k\right)\right\} \\
&= \mathbb{E}\left\{\mathcal{A}rg\left(Ne^{j\theta} + \sum_{k=0}^{N-1} n_k s_k\right)\right\} \\
&= \mathbb{E}\left\{\mathcal{A}rg\left(e^{j\theta} + \frac{1}{N}\sum_{k=0}^{N-1} n_k s_k\right)\right\} \\
&= \mathbb{E}\left\{\theta + \mathcal{A}rg\left(1 + e^{-j\theta}\sum_{k=0}^{N-1} n_k s_k\right)\right\} \\
&= \theta
\end{aligned}
$$

The last step has the following justification: The term in the summation has the expected magnitude of zero. Therefore, the quantity inside the $\mathcal{A}rg$ function is a complex Gaussian random variable with mean magnitude 1 and mean phase 0, because it is circularly symmetric.

The estimator is therefore unbiased. Estimation of the mean of estimator performed via a sample size 100 is given in Fig. 4. It also shows that the estimator is unbiased. Calculating the the variance of estimation error then needs the following evaluation:

$$
\mathbb{E}\left\{\left(\mathcal{A}rg\left(w_n\right)\right)^2\right\}
$$

where $w_n = 1 + e^{-j\theta}\sum_{k=0}^{N-1} n_k s_k$. However it is unnecessarily cumbersome. We compute the variance by simulation. The results is shown in Fig. 5

iii. We now determine the CRLB for the mean square estimation error in unknown parameter θ. From the numerical results of part (ii), we find that the estimator is unbiased. Therefore, the numerator part of the

Information Inequality is 1. We calculate the denominator, that is, $I_{\mathbf{r}}(\theta)$. We make use of (10) here.

$$
\begin{aligned}
I_{\mathbf{r}}(\theta) &= -\mathbb{E}\left\{\frac{\partial^2}{\partial\theta^2}\ln p(\mathbf{r}|\theta)\right\} \\
&= \mathbb{E}\left\{\frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Re\{r_k^* s_k e^{j\theta}\}\right\} \\
&= \mathbb{E}\left\{\frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Re\{(s_k^* e^{-j\theta} + n_k^*) s_k e^{j\theta}\}\right\} \\
&= \mathbb{E}\left\{\frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Re\{1 + n_k^* s_k e^{j\theta}\}\right\} \\
&= \frac{2}{\sigma^2}\sum_{k=0}^{N-1} 1 + \frac{2}{\sigma^2}\sum_{k=0}^{N-1}\Re\{s_k \mathbb{E}\left\{n_k^*\right\} e^{j\theta}\} \\
&= \frac{2N}{\sigma^2} + 0 \\
&= \frac{2N}{\sigma^2}
\end{aligned}
$$

Hence, the CRLB for the variance of estimation error is $\frac{\sigma^2}{2N}$.

iv. For $N = 20$, the plot of CRLB vs $\frac{1}{\sigma^2}$ is given below in Fig. 5. The variance of estimation error calculated with a sample size of 100 is also plotted.

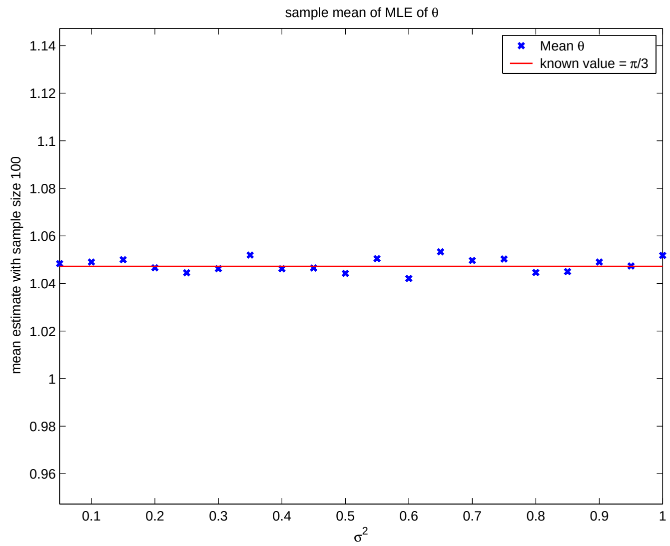

Figure 4: The plot of sample mean of the estimator at different variance values, sample size is 100 in each case.

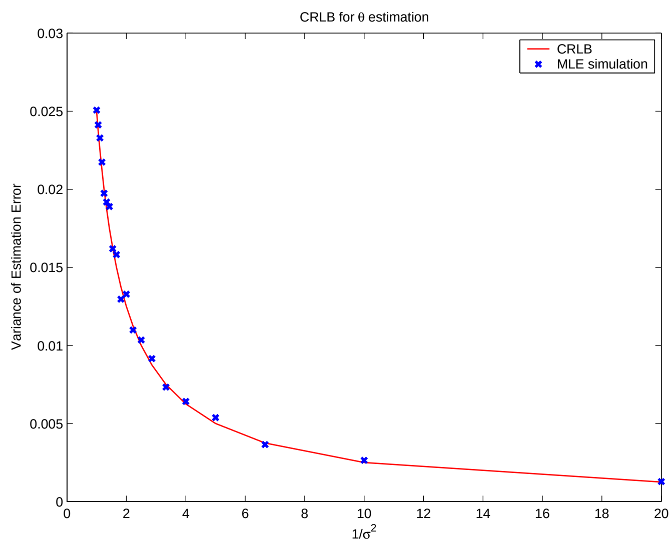

Figure 5: The plot of CRLB and simulated estimator's error variance as a function of noise variance $\frac{1}{\sigma^2}$.