Group 61
Modelling and simulation

EES101 Assignment 2

Razan Ayed
Hariharan Gopinath

October 25, 2020

1 Introduction

- In this section, we will discuss the unbiased estimator in the first part of first problem and in the second part, the maximum likelihood and the least square in the second part of the first problem.
- In 2nd problem, the system identification of the linear models will be compared and analysed.

2 First part - Estimators, Maximum Likelihood and Least Squares.

2.1 Unbiased Estimate
Given the data sample x[0],...,x[N-1] where the each sample is uniformly distributed for the interval $[0, \theta]$
That means, the range of the $\theta = 0 < \theta < \infty$
The mean for the uniform distribution can be found by mean formula
For the interval u[a,b], the mean= (a+b)/N, where N-no of data points
so, the mean value for our data points will be like

$$\frac{1}{N}\sum_{N-1}^{i=0} x(i) = \frac{0+\theta}{2} \tag{1}$$

Where replacing Ex(i)=x(i) and $\theta = \hat{\theta} inequation(1)$
Rearranging the above expression,

$$\hat{\theta} = \frac{2}{N}\sum_{N-1}^{i=0} x(i) \tag{2}$$

where it become the estimate of θ
Taking the estimate on both the side, then the expression becomes,

$$E[\hat{\theta}] = \frac{2}{N}\sum_{N-1}^{i=0} E[x(i)] \tag{3}$$

Further the equation can be simplified as,

$$E[\hat{\theta}] = \frac{2}{N}.\frac{\theta}{2} * N \tag{4}$$

$$E[\hat{\theta}] = \theta \tag{5}$$

From the equation (5), we can conclude that, this is the unbiased estimator for θ

2.2 Sample mean and Maximum Likelihood

Given the data sample

$$x[k] = A + W[k] \tag{6}$$

where k=0,1,....N-1
The W[k] begin a white Gaussian noise.
Further, the parameter A sample mean is given by

$$\hat{A} = \frac{1}{N}\sum_{k=0}^{N-1} x(i) \tag{7}$$

Here, our task is to Show that the sample mean is a Maximum Likelihood estimator (MLE) of the parameter A.
The MLE of any probability distribution function can be written as

$$L(A) = \mathbb{P}[W(K)]$$

$$L(A) = \prod_{k=0}^{N-1} \frac{1}{\sqrt{(2*\pi)}\sigma_k} * e^{\frac{-1}{2*\sigma_k^2}(x_k - A)^2)} \tag{8}$$

In order to apply for MLE, we need to maximize the the equation(8) with θ

$$\Delta_A \mathbb{P}[W(K)] = \sum_{i=0}^{N-1}\prod_{k=0}^{N-1} \frac{1}{\sqrt{(2*\pi)}\sigma_k} * e^{\frac{-1}{2*\sigma_k^2}(x_k - A)^2}(\frac{1}{2*\sigma_i^2}(x_i - A)) \tag{9}$$

$$= \mathbb{P}[W(K)]\sum_{i=0}^{N-1} \frac{1}{2\sigma_i^2}(x_i - A) \tag{10}$$

$\Delta_A \mathbb{P}[W(K)]$ =0 is achieved by

$$\sum_{i=0}^{N-1} \frac{1}{2\sigma_i^2}(x_i - A) = 0 \tag{11}$$

$$\sum_{i=0}^{N-1}(x_i) = \sum_{i=0}^{N-1}(A) \tag{12}$$

$$\sum_{i=0}^{N-1}(x_i) = N.(A) \tag{13}$$

Which can be seen is equal to the mean of the set of samples x(i)

$$\hat{A} = \frac{1}{N}\sum_{i=0}^{N-1}(x_i) \tag{14}$$

2.3(a) Linear least squares

$$\hat{\theta_N} = (R_N)^{-1} f_N \tag{15}$$

where the value of R_N and f_N is mentioned below

$$R_N = \frac{1}{N}\sum_{k=0}^{N-1} \varphi(i)\varphi^T(i)$$
$$f_N = \frac{1}{N}\sum_{i=0}^{N-1} \varphi(i)y(i)$$

The output data generated will be in the form of $y(i) = \theta^T\varphi(i) + e(i)$ where $e(i)$ is independent identically distributed(iid) random variable.
By subtituting the values of R_N and f_N, $\varphi(i) = u(k)$, in $\hat{\theta_N}$, we get the parameterised value.

$$\hat{\theta} = \left(\frac{1}{N}\sum_{k=0}^{N-1} u(k)u^T(k)\right)^{-1}\left(\frac{1}{N}\sum_{k=0}^{N-1} u(k)y(k)\right)$$
$$\hat{\theta} = \frac{\left(\frac{1}{N}\sum_{k=0}^{N-1} u(k)y(k)\right)}{\left(\frac{1}{N}\sum_{k=0}^{N-1} u(k)^2\right)}$$

2.3(b) Now, we are taking the estimation for the parameterised vector to get the biased value,

$$E[\hat{\theta}] = E\left(\frac{\left(\frac{1}{N}\sum_{k=0}^{N-1} u(k)y(k)\right)}{\left(\frac{1}{N}\sum_{k=0}^{N-1} u^2(k)\right)}\right)$$

Now by substituting the value of $y(k)$, we get the below expression.

$$E[\hat{\theta}] = E(\frac{(\frac{1}{N}\sum_{k=0}^{N-1} \theta u(k)^2 + e[k]u[k])}{(\frac{1}{N}\sum_{k=0}^{N-1} u^2(k))})$$
$$E[\hat{\theta}] = E(\theta) + E(\frac{(\frac{1}{N}\sum_{k=0}^{N-1} e[k]u[k])}{(\frac{1}{N}\sum_{k=0}^{N-1} u^2(k))})$$
$$E[\hat{\theta}] = \theta + E(\frac{(\frac{1}{N}\sum_{k=0}^{N-1} e[k]u[k])}{(\frac{1}{N}\sum_{k=0}^{N-1} u^2(k))})$$

Thus this shows that the estimation is biased.
2.3(C)
If $u[k] = 0, \forall k$, then $\hat{\theta}$ becomes zero.
2.4 In-variance property of the maximum likelihood estimate
Consider the data samples

$$y[k] = A + e[k] \tag{16}$$

Where k=0,1,.....N-1 e[k] being the Gaussian white noise with variance σ^2
The y have also given an estimate for a transformed parameter

$$\alpha = e^A \tag{17}$$

The question is to show that the maximum likelihood estimate of α is

$$\hat{\alpha} = e^{\hat{A}} \tag{18}$$

Taking the ln on both the side of the equation(21)

$$\ln \alpha = \ln e^{A} \tag{19}$$

further, it can be simplified as

$$\ln \alpha = A \ln_e e \tag{20}$$

$$\ln \alpha = A \tag{21}$$

The maximum likelihood of A can be calculated as $\hat{A}$

$$\hat{A} = \frac{1}{N} \sum_{i=0}^{N-1} y(t) \tag{22}$$

subs the equation (25) in the equation(20)

$$y[k] = \ln \alpha + e[k] \tag{23}$$

Apply the maximum likelihood on α,

$$L(\alpha)\text{=}\mathbb{P}[\text{e(K)}]$$

$$L(\alpha) = \prod_{k=0}^{N-1} \frac{1}{\sqrt{(2\pi)}\sigma_k} * e^{\frac{-1}{2\sigma_k^2}(y_k - \ln \alpha)^2} \tag{24}$$

To maximize over the α,

$$\bigtriangledown_\theta \mathbb{P}[W(K)] = \sum_{i=0}^{N-1} \prod_{k=0}^{N-1} \frac{1}{\sqrt{(2\pi)}\sigma_k} * e^{\frac{-1}{2\sigma_k^2}(y_k - \ln \alpha)^2} (\frac{1}{2\sigma_i^2}(y_k - \ln \alpha)) \tag{25}$$

From the equation(28), we can simplify the above equation as,

$$\bigtriangledown_\theta \mathbb{P}[W(K)] = \mathbb{P}[e(K)] \sum_{i=0}^{N-1} (\frac{1}{2\sigma_i^2}(y_k - \ln \alpha)) \tag{26}$$

$\bigtriangledown_\theta \mathbb{P}[W(K)] = 0$ only when

$$\sum_{i=0}^{N-1} (\frac{1}{2\sigma_i^2}(y_k - \ln \alpha)) = 0 \tag{27}$$

Taking the sum inside the equation,

$$\sum_{i=0}^{N-1} y_i = \sum_{i=0}^{N-1} \ln \alpha \tag{28}$$

Further, this equation can be simplified as

$$\sum_{i=0}^{N-1} y_i = N \ln \alpha \tag{29}$$

$$\frac{1}{N} \sum_{i=0}^{N-1} y_i = \ln \alpha \tag{30}$$

Taking exponential on both side of the equation(33) therefore we can derive for the $\hat{\alpha}$,

$$e^{\frac{1}{N}\sum_{i=0}^{N-1} y_i} = \hat{\alpha} \tag{31}$$

We know that

$$\hat{A} = \frac{1}{N}\sum_{i=0}^{N-1} y(t)$$

Therefore,

$$\hat{\alpha} = e^{\hat{A}}$$

3 Identification of linear models for dynamical systems

3.1 One step ahead predicator

3.1.1 a

The model structure we got is

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + e(t) + c_1 e(t-1)$$

We had to find out which kind of model structure is it.
In order to determine that we had to use the Z-transform as following:

$$\begin{aligned} y(k) + a_1 Y(k)Z^{-1} + a_2 Y(k)z^{-2} = b_0 U(k) + E(k) + c_1 E(k)Z^{-1} \\ Y(k) = \frac{b_0 U(k) + E(k) + c_1 E(k)Z^{-1}}{1 + a_1 Z^{-1} + a_2 Z^{-2}} \\ Y(k) = \frac{b_0}{1 + a_1 Z^{-1} + a_2 Z^{-2}} U(k) + \frac{1 + c_1 Z^{-1}}{1 + a_1 Z^{-1} + a_2 Z^{-2}} E(k) \end{aligned}$$

. From the past equation we could see that it's an ARMAX model structure because it's similar to the equation

$$y = (\frac{A}{B})u + (\frac{C}{A})e \tag{32}$$

3.1.2 b

Here we had to find out how does the plant model G and the noise model H looks like.
From equation 35 we could see that

$$G = \frac{b_0}{1 + a_1 Z^{-1} + a_2 Z^{-2}} \tag{33}$$

and

$$H = \frac{1 + c_1 Z^{-1}}{1 + a_1 Z^{-1} + a_2 Z^{-2}} \tag{34}$$

3.1.3 c

In order to find the 1-step-ahead predicator for the model, we need to use the equation:

$$\hat{y}(t|t-1,\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t)$$
$$\hat{y}(t|t-1,\theta) = \frac{(1 + a_1Z^{-1} + a_2Z^{-2})b_0}{(1 + c_1Z^{-1})(1 + a_1Z^{-1} + a_2Z^{-2})}u(t) + \frac{C_1Z^{-1} - a_1z^{-1} - a_2z^{-2}}{1 + c_1z^{-1}}y(t)$$
$$(1 + c_1z^{-1})\hat{y}(t|t-1,\theta) = b_0u(t) + (c_1z^{-1} - a_1z^{-1} - a_2z^{-2})y(t)$$
$$\hat{y}(t|t-1,\theta) + c_1z^{-1}\hat{y}(t|t-1,\theta) = b_0u(t) + c_1z^{-1}y(t) - a_1y(t-1) - a_2y(t-2)$$
$$\hat{y}(t|t-1,\theta) = \frac{b_0u(t) + c_1z^{-1}y(t) - a_1y(t-1) - a_2y(t-2)}{(1 + c_1z^{-1})}$$

3.1.4 d

Nonlinear

3.2 2

prediction or simulation
(a) Consider the following model structure

$$y(t) + a_1y(t-1) = b_0u(t) + e(t) + a_1e(t-1) \tag{35}$$

We had to find out which kind of model structure is it.
In order to determine that we had to use the Z-transform as following:

$$y(t) + a_1y(k)Z^{-1} = b_0u(t) + e(t) + a_1e(k)Z^{-1} \tag{36}$$

$$y(t) = \frac{b_0u(k) + e(k) + a_1e(k)Z^{-1}}{a_1y(k)Z^{-1}} \tag{37}$$

$$y(t) = \frac{b_0}{1 + a_1Z^{-1}}u(k) + \frac{1 + a_1Z^{-1}}{1 + a_1Z^{-1}}e(k)$$

This is an OD kind of model structure
(b) Find the 1-step-ahead predictor for the model

$$\hat{y}(t|t-1,\theta) = H^{-1}(q,\theta)G(q,\theta)u(t) + (1 - H^{-1}(q,\theta))y(t)$$

$$\hat{y}(t|t-1,\theta) = \frac{b_0}{1 + a_1Z^{-1}}u(t) + 0$$

$$\hat{y}(t|t-1,\theta) * (1 + a_1Z^{-1}) = b_0u(t)$$

$$\hat{y}(t|t-1,\theta) + (\hat{y}(t|t-1,\theta) * a_1Z^{-1}) = b_0u(t)$$

$$\hat{y}(t|t-1,\theta) + (\hat{y}(t-1|t-2,\theta) * a_1) = b_0u(t)$$

$$\hat{y}(t|t-1,\theta) = b_0u(t) - a_1(\hat{y}(t-1|t-2,\theta))$$

This is not the previous output and also the old values are not present in the value. So, this looks like simulation function.

3.3 3

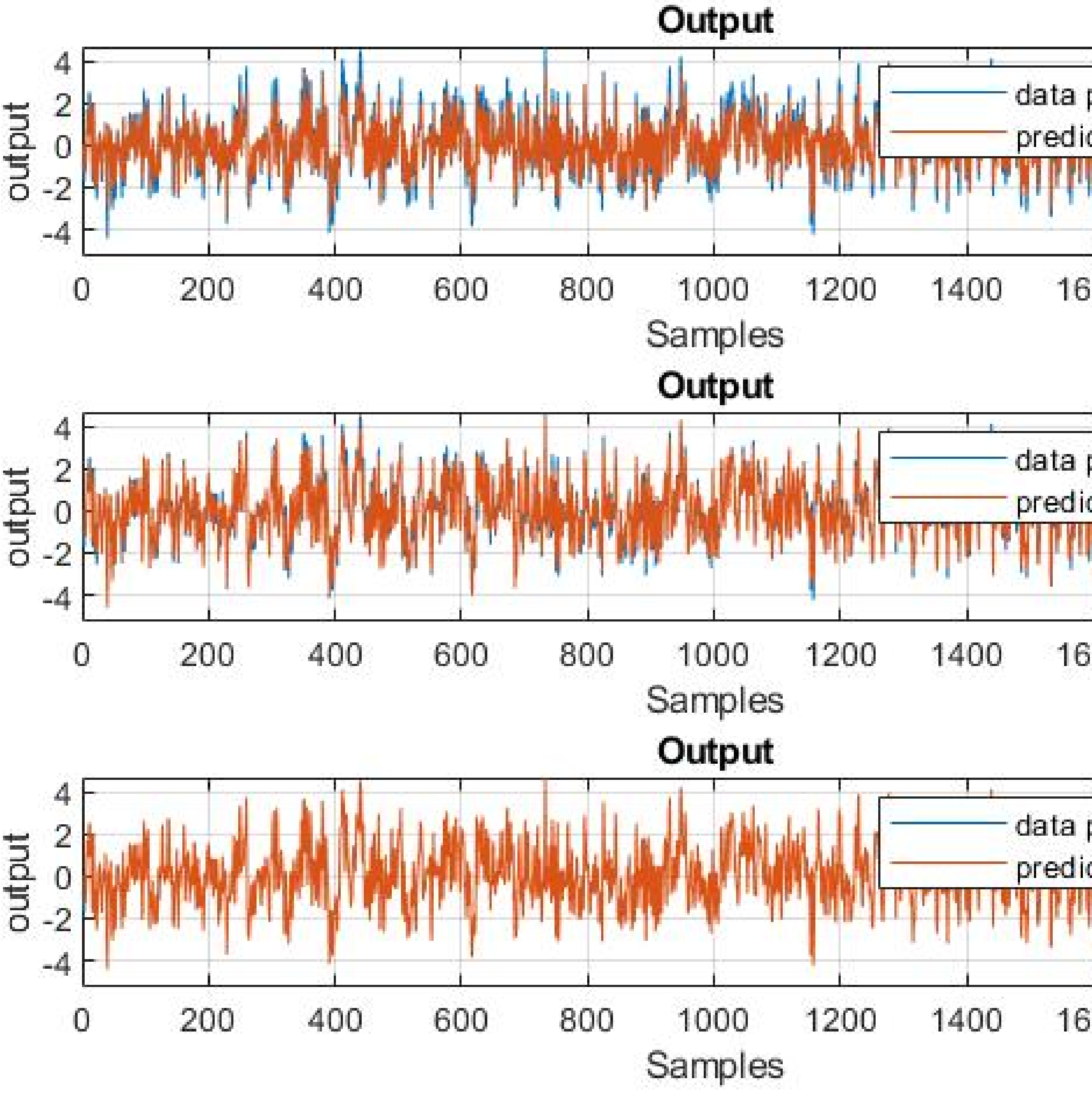
Here we had three different candidate ARX model structures which are:

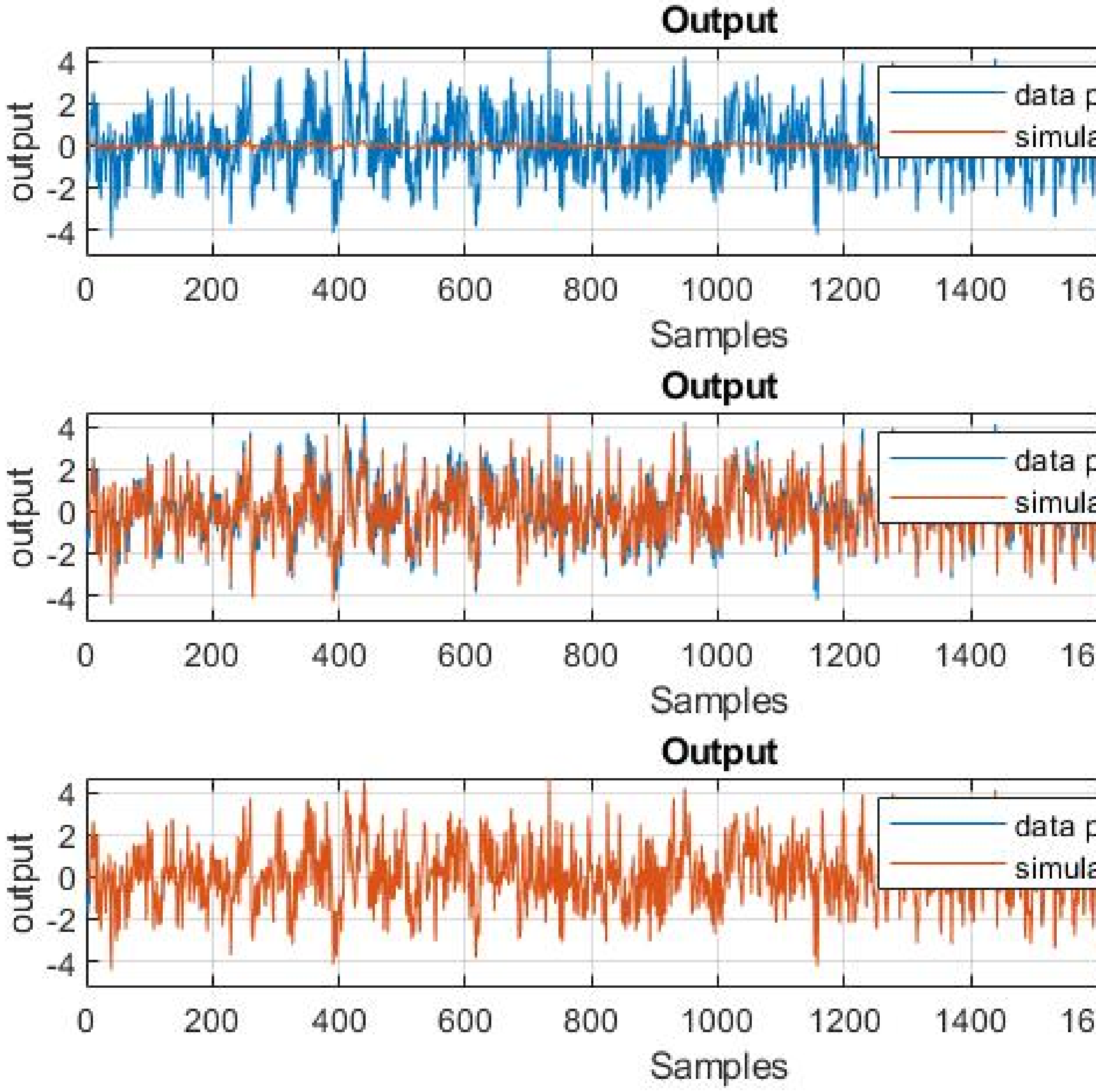
$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + e(t)$$
$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + b_1 u(t-1) + e(t)$$
$$y(t) + a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3) = b_1 u(t-1) + e(t)$$

The question was about deciding which model of the past three is the best one of the system.
In order to decide which one is the best we had to find out which one of them provides the minimum error.
First we used the 1-step ahead predictor to define the linear regression in the parameters.
As a result we got the following root mean squares for each model:

Model nr	simRMS	predRMS
1	1.3950	1.0587
2	0.4231	0.3200
3	0.0591	0.0493

As a result we can see that the third model is the best one as it provides the least error.


Output
output
4
2
0
-2
-4
0
200
400
600
800
1000
1200
1400
Samples
Output
output
Samples
Output
output
Samples


Output
output
data p
simula
Samples
Output
output
data p
simula
Samples
Output
output
data p
simula
Samples