Modelling and Simulation

ES5101 - Home Assignment 2

Varun Hegde   Arshad N

October 4, 2019

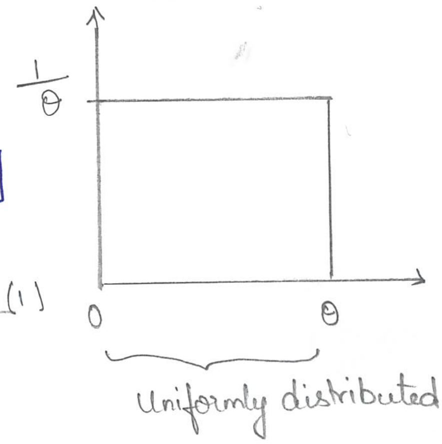
1. Introduction

In this assignment we would be working with two problems. In problem 1 Unbiased estimator will be used and in second part of problem 1, we used Maximum likelihood and least square. In problem 2 System Identification of Linear model will be analyzed and compared.

2. First part - Estimators, Maximum likelihood & least Square

2.1 Unbiased estimate

a) Given data samples $x[0], \ldots, x[N-1]$ distributed $\mathcal{U}[0, \theta]$



The range of $\theta = 0 < \theta < \infty$

The mean for Uniform distribution $\mathcal{U}[a, b]$

$$\text{Mean} = \frac{a+b}{2} \qquad (1)$$

likethat, mean of data samples

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

where $E x(i) = x(i)$ and $\theta = \hat{\theta}$

Rearranging, $$\hat{\theta} = \frac{2}{N}\sum_{i=0}^{N-1} x(i) \qquad (3)$$

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

$$E[\hat{\theta}] = \theta \qquad (4)$$

This is an unbiased estimator for θ.

2.2 Sample mean and Maximum likelihood

a) Given data samples

$$x[k] = A + w[k] \quad \text{where } k = 0 \ldots\ldots N-1$$

Then the sample mean $\hat{A} = \frac{1}{N}\sum_{k=0}^{N-1} x[k]$ ——— (5).

Maximum likelihood function

$$L(A) = \mathbb{P}\,[wk] \qquad (6)$$

$$= \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi}\,\sigma_k} e^{-\frac{1}{2\sigma_k^2}(x_k - A)^2} \qquad (7)$$

In order to apply the MLE, we need to maximize (7) over θ

Therefore, $$\nabla_A \mathbb{P}\,[w(k)] = \sum_{i=0}^{N-1} \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi}\,\sigma_k} e^{-\frac{1}{2\sigma_k^2}(x_k-\theta)^2} \frac{1}{\sigma_i^2}(x_i - A) \qquad (8)$$

$$= \mathbb{P}\,w[k] \cdot \sum_{i=0}^{N-1} \frac{1}{\sigma_i^2}(x_i - A) \qquad (9)$$

$\nabla_\theta \mathbb{P}\,[w(k)] = 0$ is achieved by

$$\sum_{i=0}^{N-1} \frac{1}{\sigma_i^2}(x_i - A) = 0 \qquad (10)$$

and yields: $$\hat{A} = \frac{1}{\sum_{i=0}^{N-1} \frac{1}{\sigma_i^2}} \sum_{i=0}^{N-1} \frac{1}{\sigma_i^2} x_i \qquad (11)$$

If all the measurements have the same variance, then (11) boils down to a simple average.

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

2.3 Linear least squares

a). Consider the Linear System

$$y[k] = \theta u[k] + e[k] \quad \text{(13)}$$

where $k = 0, \dots, N-1$

Then $$e(k) = y(k) - \theta u(k) \quad \text{(14)}$$

least square estimate $$\hat{\theta} = \arg\max_{\theta} \mathbb{P}\left[e = y(k) - \theta u(k)\right] \quad \text{(15)}$$

$$P[e = x] = \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2\sigma_k^2}\left(y(k) - \theta u(k)\right)^2} \quad \text{(16)}$$

$$\hat{\theta} = \arg\max_{\theta} \prod_{k=0}^{N-1} e^{-\frac{1}{2}\left(\frac{y - \theta u}{\sigma^2}\right)^2} \quad \text{(17)}$$

both of n and its log of fn achieve max at same point

$$\hat{\theta} = \arg\max \log \prod_{k=0}^{N-1} e^{-\frac{1}{2}\left(\frac{y(k) - \theta u(k)}{\sigma^2}\right)^2}$$

$$\hat{\theta} = \arg\max_{\theta} -\left[\sum_{k=0}^{N-1} \frac{1}{2} \frac{(y(k) - \theta u(k))^2}{\sigma^2}\right] \quad \text{(18)}$$

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

solution to this least square estimate

$$\hat{\theta} = \min_{\theta} \sum_{k=0}^{N-1} \frac{1}{2} \frac{(y(k) - \theta u(k))^2}{\sigma^2} \quad \text{(19)}$$

solution implies obtaining roots of gradient fn

let $$J(\theta, y) = \sum_{k=0}^{N-1} \frac{1}{2} \left\| \frac{\boxed{y - \theta u(k)}}{\sigma^2} \right\|^2 \quad \text{(20)}$$

$$\nabla_\theta J(\theta, y) = \nabla_\theta \sum_{k=0}^{N-1} \frac{1}{2} \frac{\| y - \theta u(k) \|^2}{\sigma^2} = 0 \tag{21}$$

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

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

$$\sum_{k=0}^{N-1} u(k)\, y(k) = \hat{\theta} \sum_{k=0}^{N-1} u(k)^2$$

$$\hat{\theta} = \frac{\sum_{k=0}^{N-1} u(k)\, y(k)}{\sum_{k=0}^{N-1} u(k)^2} \tag{22}$$

b) From Unbiased Condition $E(\hat{\theta}) - \theta^* = 0$.

Also $y(k) = \theta^* u(k) + e(k)$

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

$$= E\left(\frac{\theta^* \sum_{k=0}^{N-1} u(k)^2}{\sum_{k=0}^{N-1} u(k)^2} + \frac{\sum_{k=0}^{N-1} e(k)\, u(k)}{\sum_{k=0}^{N-1} u(k)^2} \right)$$

$$E[\hat{\theta}] = \theta^* + \frac{1}{\sum_{k=0}^{N-1} u(k)^2} E\left(\sum_{k=0}^{N-1} e(k)\, u(k) \right) \tag{23}$$

$\boxed{E[\theta^*] = \theta^*}$

From ( ) we conclude $E[\hat{\theta}] - \theta \neq 0$, It is biased

c) if $u(k) = 0$, then the Estimation $E[\hat{\theta}]$ also becomes 0.

$$E[\hat{\theta}] = 0 \tag{24}$$

2.4 Invariance property of the maximum likelihood

Consider the data sample,

$$y[k] = A + e[k] \qquad (25)$$

where $k = 0 \ldots N-1$

with $e[k]$ being WGN with Variance σ^2

Also given $\quad \alpha = e^{A} \qquad (26)$

To show $\quad \hat{\alpha} = e^{\hat{A}}$

Let $\quad \alpha = e^{A}$

$$\ln \alpha = \ln e^{A} \qquad (27)$$

$$= A \ln_e e$$

$$\ln \alpha = A \quad \text{and} \quad \hat{A} = \frac{1}{N} \sum_{i=0}^{N-1} y(k)$$

$$y[k] = \ln \alpha + e[k] \qquad (28)$$

Application of MLE on α

$$L(\alpha) = \mathbb{P}\left[e[k]\right]$$

$$= \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2\sigma_k^2}\left[y_k - \ln\alpha\right]^2} \qquad (29)$$

To maximize over α,

$$\nabla_\theta \mathbb{P}\left[e(k)\right] = \sum_{i=0}^{N-1} \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2\sigma_k^2}\left[y_k - \ln\alpha\right]^2} \cdot \frac{1}{\sigma_i}(y_k - \ln\alpha$$

$$= \mathbb{P}\left[e(k)\right] . \sum_{i=0}^{N-1} \frac{1}{\sigma_i^2} (y_i \ln\alpha) \qquad (30)$$

$\nabla_\theta \mathbb{P}\left[e(k)\right] = 0$ only when

$$\sum_{i=0}^{N-1} \frac{1}{\ldots} (y_i - \ln\alpha) = 0 \qquad (31)$$

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

$$A = \sum_{i=1}^{N-1} \ln \alpha \quad \text{(32)}$$

Therefore, $e^{\hat{A}} = \sum_{i=1}^{N-1} \alpha$

$$= \hat{\alpha}$$

which is nothing but $\hat{\alpha} = e^{\hat{A}}$ (33)

3. Second part - Identification of linear Models for dynamical Systems

3.1 One-Step ahead predictor

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + e(t) + c_1 e(t-1) \quad \text{(34)}$$

a) $y(t) + a_1 q^{-1} y(t) + a_2 q^{-2} y(t) = b_0 u(t) + e(t) + c_1 q^{-1} e(t)$

$$(1 + a_1 q^{-1} + a_2 q^{-2})\, y(t) = b_0 u(t) + (1 + c_1 q^{-1})\, e(t)$$

$$A(q)\, y(t) = B(q)\, u(t) + C(q)\, e(t)$$

$$\boxed{y(t) = \frac{B(q)}{A(q)} u(t) + \frac{C(q)}{A(q)} e(t)} \quad \text{(35)}$$

This is ARMAX model structure

b) Expression of the plant model G and nois model H

$$G(q) = \frac{B(q)}{A(q)}$$

$$G(q) = \frac{b_0}{(1 + a_1 q^{-1} + a_2 q^{-2})} \quad (36)$$

Then,
$$H(q) = \frac{C(q)}{A(q)}$$

$$H(q) = \frac{1 + c_1 q^{-1}}{(1 + a_1 q^{-1} + a_2 q^{-2})} \quad (37)$$

$$\theta = \begin{bmatrix} a_1 \\ a_2 \\ b_0 \\ c_1 \end{bmatrix}$$

c) 1-step ahead predictor for Model

$$\hat{y}(t|t-1) = H^{-1} G\, u(t) + (1 - H^{-1})\, y(t) \quad (38)$$

$$\hat{y}(t|t-1) = \left(\frac{1 + a_1 q^{-1} + a_2 q^{-2}}{1 + c_1 q^{-1}}\right)\left(\frac{b_0}{1 + a_1 q^{-1} + a_2 q^{-2}}\right) u(t)$$

$$+ \left(\frac{1 + c_1 q^{-1} - 1 - a_1 q^{-1} - a_2 q^{-2}}{1 + c_1 q^{-1}}\right) y(t)$$

$$\hat{y}(t|t-1) = \frac{b_0}{(1 + c_1 q^{-1})} u(t) + \left(\frac{c_1 q^{-1} - a_1 q^{-1} - a_2 q^{-2}}{1 + c_1 q^{-1}}\right) y(t)$$

$$(1 + c_1 q^{-1})\, \hat{y}(t|t-1) = b_0 u(t) + (c_1 q^{-1} - a_1 q^{-1} - a_2 q^{-2})\, y(t)$$

$$\hat{y}(t|t-1) + c_1\hat{y}(t-1|t-2) = b_0u(t) + (c_1q^{-1} - a_1q^{-1} - a_2q^{-2})y(t)$$

$$\hat{y}(t|t-1) = b_0u(t) - a_1y(t-1) - a_2y(t-2) + c_1(y(t-1) - \hat{y}(t-1|t-2)) \quad (39)$$

d) No, This is a non linear function of parameter because old predictor values are present in the 1-step predictor

3.2 Prediction or Simulation

$$y(t) + a_1y(t-1) = b_0u(t) + e(t) + a_1e(t-1) \quad (40)$$

a)

$$y(t) + a_1q^{-1}y(t) = b_0u(t) + e(t) + a_1q^{-1}e(t)$$

$$(1+a_1q^{-1})y(t) = b_0u(t) + (1+a_1q^{-1})e(t)$$

$$A(q)y(t) = B(q)u(t) + A(q)e(t)$$

$$y(t) = \frac{B(q)}{A(q)}u(t) + \frac{A(q)}{A(q)}e(t)$$

$$y(t) = \frac{B(q)}{A(q)}u(t) + e(t) \quad (41)$$

This is OE Kind of Model Structure

b) $\hat{y}(t|t-1) = H^{-1} G\, u(t) + (1 - H^{-1})\, y(t)$ ______ (42)

$= 1 \times \left(\frac{B(q)}{A(q)}\right) u(t) + \cancel{(1-1)}\, y(t)$ $\rightarrow 0$

$= \frac{B(q)}{A(q)}\, u(t)$

$= \frac{b_0}{(1 + a_1 q^{-1})}\, u(t)$

$\hat{y}(t|t-1)(1 + a_1 q^{-1}) = b_0 u(t)$

$\hat{y}(t) = -a_1 \hat{y}(t-1) + b_0 u(t)$

$\hat{y}(t) = b_0 u(t) - a_1 \hat{y}(t-1)$ ______ (43)

This is not the previous output also old values are not present in this function. This looks like a simulation function.

3.3 Identification of an ARX model

3.3.a) $y(t) + a_1 y(t-1) + a_2 y(t-2) = b_0 u(t) + e(t)$

$$\hat{y} = \left[u(t) \;\; -y(t-1) \;\; -y(t-2) \right] \begin{bmatrix} b_0 \\ a_1 \\ a_2 \end{bmatrix} = \varphi^T \theta \quad \text{______ (44)}$$

θ will be found to minimize the prediction error for the given

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{N} \Sigma(\varepsilon^T \varepsilon) = \arg\min_{\theta} \sum_{t=1}^{N} (y_t - \varphi_t^T \theta)^2$$

$$= \arg\min_{\theta} (Y - \phi^T\theta)^T (Y - \phi^T\theta) \quad (45)$$

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} (y_t - \varphi_t^T \theta)^2 = (y_1 - \varphi_1^T\theta)^2 + (y_2 - \varphi_2^T\theta)^2 + \dots + (y_N - \varphi_N^T\theta)^2$$

$$= [(y - \varphi_1^T\theta)(y_2 - \varphi^T\theta)\dots] \begin{bmatrix} y_1 - \varphi_1^T\theta \\ \vdots \\ y_N - \varphi_N^T\theta \end{bmatrix}$$

$$= (Y - \phi^T\theta)^T (Y - \phi^T\theta) \quad (46)$$

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

$$y(t) + a_1 q^{-1} y(t) + a_2 q^{-2} y(t) = b_0 u(t) + e(t)$$

$$y(t)(1 + a_1 q^{-1} + a_2 q^{-2}) = b_0 u(t) + e(t)$$

$$A = [1 + a_1 q^{-1} + a_2 q^{-2}] \qquad B = b_0 \quad (47)$$

b) Simillarly, $A = [1 + a_1 q^{-1} + a_2 q^{-2}]$ $B = [b_0 + b_1 q^{-1}]$ (48)

$$\therefore \quad b_1 u(t-1) = b_1 q^{-1}$$

c) Simillarly, $A = [1 + a_1 q^{-1} + a_2 q^{-2} + a_3 q^{-3}]$ $B = [b_1 q^{-1}]$

(49)

One step predictor for the Models

a) $\hat{y}(t|t-1) = B(q,\theta)\,u(t) + (1 - A(q,\theta))\,y(t)$

$= b_0 u(t) + [-(a_1 y(t-1)) + a_2 y(t-2)]$ ______ (50)

b) $\hat{y}(t|t-1) = b_0 u(t) + b_0 u(t-1) - [a_1 y(t-1) + a_2 y(t-2)]$ ______ (51)

c) $\hat{y}(t|t-1) = b_1 u(t-1) - [a_1 y(t-1) + a_2 y(t-2) + a_3 y(t-3)]$ ______ (52)

Then,

The parameters of 1st model are

$b_0 = 0.0688$ $\quad a_1 = -0.9590$ $\quad a_2 = 0.367$ ______ (53)

The parameters of 2nd model are

$b_0 = 0.0113$ $\quad b_1 = 0.9946$ $\quad a_1 = -0.8923$ $\quad a_2 = 0.3116$ ______ (54)

The parameters of 3rd model are

$b_1 = 1$ $\quad a_1 = -1$ $\quad a_2 = 0.6$ $\quad a_3 = -0.3$ ______ (55)

3.3. b)

The RMSE Values of prediction models are

Model1: 1.0584 Model2: 0.3173 Model3: 0.0096 ——— (56)

The RMSE Values of Simulation models are

Model 1: 1.3943 Model2: 0.4199 Model3: 0.0223 ——— (57)

From this we conclude that Model 3 is best it has low RMSE Values on both models.

3.3 c)

The Covariance of the parameters is defined as

$$\mathrm{Cov}(\hat{\theta}) = E\left[(\hat{\theta} - E(\hat{\theta}))(\hat{\theta} - E(\hat{\theta})^T)\right] = \sigma^2(\phi\phi^T)^{-1} \quad (58)$$

The Variance of noise is given as 0.01. Then the diagonal elements of the Covariance matrix of each model is compared according to Covariance analysis and found that Model 1 has lowest Values. Hence Model 1 is best.