Modelling and simulation

Assignment 2

ESS101

Group 16

27 september 2017

1. Part 1 Task 1

Due to the characteristic of an uniformed distribution between 0 and a given value. We find that the mean will always be in between these two, i.e. mean of x_i is $\theta/2$. Therefore we can create the estimator for θ to depend on the mean of x_i ($\bar{x}$).

$$\hat{\theta} = \bar{x} \cdot 2 \tag{1.1}$$

2. Task 2

Given a normal, centered noise on output signal of the system it was shown during the lecture that we can use least-squares instead of maximum-likelihood. Instead of maximizing the likehood of getting $\hat{A}$ as close to A as possible, we're minimizing the error between $\hat{A}$ and A. We then create the estimator of A with least-squares.

$$\hat{A} = \arg\min_{A} \sum_{k=0}^{N-1} (X_k - A)^2 \tag{2.1}$$

So to determine the estimator, we want to minimize the sum above. In order to do this we create a function f, shown below.

$$f(A) = \sum_{k=0}^{N-1} (X_k - A)^2 = \sum_{k=0}^{N-1} X_k^2 - 2A \cdot \sum_{k=0}^{N-1} X_k + NA^2 \tag{2.2}$$

In order to find our minimum, we try to find the value of A which makes the derivative of f equal zero.

$$f'(A) = -2\sum_{k=1}^{N} X_k + 2kA \tag{2.3}$$

$$f'(A) = 0 \tag{2.4}$$

This critera is only satisfied if

$$A = \frac{1}{N}\sum_{k=0}^{N-1} X_k \tag{2.5}$$

which gives us

$$\hat{A} = \frac{1}{N}\sum_{k=0}^{N-1} x[k] \tag{2.6}$$

Thus 2.6 is an estimator of A, using maximum-likelihood through least-squares.

3. Task 3

a

When we estimate θ we try to find the relation between y and u, i.e $\hat{\theta} = \frac{y[k]}{u[k]}$. If we then assume that $u[k] = 0, \forall k$, our $\hat{\theta} \to \infty$.

b

Again, we try to find the least-squares of the function.

$$\hat{\theta} = \arg\min_{\theta} \sum_{k=0}^{N-1} (\frac{y[k]}{u[k]} - \theta)^2 \tag{3.1}$$

We define our function f as following.

$$f(\theta) = \sum_{k=0}^{N-1} (\frac{y[k]}{u[k]} - \theta)^2 = \sum_{k=0}^{N-1} (\frac{y[k]}{u[k]})^2 - 2\theta \cdot \sum_{k=0}^{N-1} \frac{y[k]}{u[k]} + N\theta^2 \tag{3.2}$$

To find the minimum of the function $f(\theta)$, we find the value of θ that matches $f'(\theta) = 0$.

$$f'(\theta) = -2 \sum_{k=0}^{N-1} \frac{y[k]}{u[k]} + 2N\theta = 0 \tag{3.3}$$

$$2N \cdot \theta = 2 \cdot \sum_{k=0}^{N-1} \frac{y[k]}{u[k]} \qquad \to \qquad \hat{\theta} = \frac{1}{N} \sum_{k=0}^{N-1} \frac{y[k]}{u[k]} \tag{3.4}$$

c

The given model is linear, thus our estimator is unbiased according to the information provided during the lectures.

4. Task 4

$$\hat{\alpha} = \arg\max_{A} e^{-\frac{1}{2}\sum_{k=0}^{N-1}(y[k]-ln(\hat{\alpha}))^2} \tag{4.1}$$

This is the same as maximizing the exponent

$$\hat{\alpha} = \arg\max_{A} -\frac{1}{2}\sum_{k=0}^{N-1}(y[k]-ln(\hat{\alpha}))^2 \tag{4.2}$$

By muliplying the expression with *-1* least-squares can be used instead.

$$\hat{\alpha} = \arg\min_{A} \frac{1}{2}\sum_{k=0}^{N-1}(y[k]-ln(\hat{\alpha}))^2 = \tag{4.3}$$

$$= \arg\min_{A} \frac{1}{2}\sum_{k=0}^{N-1}(y^2 - 2y \cdot ln(\hat{\alpha}) + ln(\hat{\alpha}))^2$$

By separating the each part of the sum we get three separate sums instead

$$= \arg\min_{A} \frac{1}{2}(\sum_{k=0}^{N-1} y^2 - \sum_{k=0}^{N-1} 2y \cdot ln(\hat{\alpha}) + \sum_{k=0}^{N-1} ln(\hat{\alpha})^2) \tag{4.4}$$

To find our $\hat{a}$, we create a function f, which we will use to express the estimation.

$$f(\hat{a}) = \frac{1}{2}(\sum_{k=0}^{N-1} y^2 - \sum_{k=0}^{N-1} 2y \cdot ln(\hat{\alpha}) + \sum_{k=0}^{N-1} ln(\hat{\alpha})^2) \tag{4.5}$$

To minimize the function, we take the derivative and try to find the value of $\hat{a}$ that makes the derivative equal zero.

$$f'(a) = 0 = \sum_{k=0}^{N-1} \frac{ln(\hat{a})}{\hat{a}} - \sum_{k=0}^{N-1} \frac{y[k]}{\hat{a}} = \sum_{k=0}^{N-1} ln(\hat{a}) - \sum_{k=0}^{N-1} y[k]$$

As $\hat{a}$ is not depending on k, we can simply extract it from the sum to only multiplicate with N.

$$ln(\hat{a}) = \frac{1}{N}\sum_{k=0}^{N-1} y[k] = \hat{A}$$

$$\Rightarrow \quad \hat{a} = e^{\frac{1}{N}\sum_{k=0}^{N-1} y[k]} = e^{\hat{A}} \tag{4.6}$$

Thus proving that the estimation $\hat{a} = e^{\hat{A}}$ is appropriate.

5. Part 2 Task 1

5.1. a

As shown in the lectures we identify our model as ARMAX. The expression shown below

$$AY = BU + C\mathcal{E} \tag{5.1}$$

is a typical ARMAX. The given function contains all these parameters.

5.2. b

We start with identifying that we have a continous model. Then we try to show our equation on the following form.

$$Y(s) = G(s) \cdot U(s) + H(s) \cdot E(s) \tag{5.2}$$

In doing this, we first take the Laplace transform on our initial form.

$$y(t) \overset{L}{\Longleftrightarrow} Y(s) \cdot (1 + a_1 e^{-s} + a_2 e^{-2s}) = U(s) \cdot (b_0) + E(s) \cdot (1 + c_1 e^{-1}) \tag{5.3}$$

$$G(s) = \frac{b_0}{1 + a_1 e^{-s} + a_2 e^{-2s}} \tag{5.4}$$

$$H(s) = \frac{1 + c_1 e^{-s}}{1 + a_1 e^{-s} + a_2 e^{-2s}} \tag{5.5}$$

5.3. c

We start with observing our noise model.

$$Y(s) = G(s)U(s) + \underbrace{H(s)E(s)}_{:=v(t)} \tag{5.6}$$

To identify v(t) we create a function describing the noise model.

$$V_k = \int_{j=0}^{\infty} h_j \cdot e_{t-j} = e_t + \int_{j=1}^{\infty} h_j \cdot e_{t-j} \tag{5.7}$$

So to form $\hat{y}$, we will need $\hat{V_k}$. To identify $\hat{V_k}$, we will first identify our e_k.

$$\mathbb{E}[v_t] = \mathbb{E}[e_t] + \int_{j=1}^{\infty} h_j \mathbb{E}[\underbrace{e_{t-j}}_{\text{known}}] = \hat{v}_t \tag{5.8}$$

As e_{k-j} is from previous measurements, it can be considered to be known.

$$\hat{v}_t = \int_{j=0}^{\infty} \underbrace{h_j e_{t-j}}_{v_t} - e_t \tag{5.9}$$

In consequence, we can now write $\hat{V_k}$ as the following expression

$$\hat{V_t} = H(s)E(s) - E(s) = (H(s) - 1)E(s) \tag{5.10}$$

Thus, our function no longer contains unknown parameters. $\hat{Y}(s)$ can now be written as

$$\hat{Y}(s) = G(s)U(s) + \hat{V}(s) = G(s)U(s) + (H(s) - 1)E(s) \tag{5.11}$$

$E(s)$ can be rewritten as

$$E(s) = H(s)^{-1}V(s) \tag{5.12}$$

thus leaving us with the complete model

$$\hat{Y}(s) = G(s)U(s) + (H(s)^{-1} - 1)E(s) \tag{5.13}$$

We can then rewrite our model as

$$E(s) = H(s)^{-1}V(s) \Rightarrow \tag{5.14}$$

$$\Rightarrow \hat{Y}(s) = G(s)U(s) + (1 - H(s)^{-1})V(s)$$

In order to make the model a bit clearer, we identify the model as

$$H(s) \underbrace{\hat{Y}(s)}_{\text{predictor}} = \underbrace{G(s)U(s)}_{\text{pure model}} + \underbrace{(H(s) - 1)Y(s)}_{\text{past measurements}} \tag{5.15}$$

5.4. d

Due to the function being recursive when computing several predictions, it will be nonlinear. We get a similar behaviour as the example below.

$$\begin{aligned}
\hat{y}_1 &= a\hat{y}_0 + bu_0 \\
\hat{y}_2 &= a^2\hat{y}_0 + abu_0 + bu_1 \\
\hat{y}_3 &= a^3\hat{y}_0 + a^2bu_0 + abu_1 + bu_2 \\
&\vdots
\end{aligned}$$

We clearly see that the model becomes non-linear with past dependencies.

6. Task 2

6.1. a

As shown in the lectures we identify our model as ARX. The expression shown below

$$Y = \frac{B}{A}U + \frac{1}{A}\mathcal{E} \tag{6.1}$$

is a typical ARX. The given function contains all these parameters.

6.2. b

We start by doing the same procedure as in Task 1c. Laplace-transforming the model gives us

$$Y(s)(1 + a_1 e^{-s}) = b_o U(s) + E(s)(1 + a_1 e^{-s}) \tag{6.2}$$

If we then start defining our G and H

$$G = \frac{b_0}{1 + a_1 e^{-s}} \qquad and \qquad H = \frac{1 + a_1 e^{-s}}{1 + a_1 e^{-s}} = 1 \tag{6.3}$$

According Task 1c we will get the following predictor for $\hat{Y}$

$$H(s) \underbrace{\hat{Y}(s)}_{\text{predictor}} = \underbrace{G(s)U(s)}_{\text{pure model}} + \underbrace{(H(s) - 1)Y(s)}_{\text{past measurements}} \tag{6.4}$$

Due to our $H(s) = 1$, the model will take the following the form.

$$\hat{Y}(s) = G(s)U(s) \tag{6.5}$$

Thus making our predictor follow the "pure model".

7. Task 3

7.1. a

We split the data similar to the computer exercise and provided MATLAB-code on the course page $SysID\_linear\_model\_ident\_BLANK.m$. We solve the estimated values for Θ with least-squares in a similar fashion as in the equations (3.1) to (3.4), but with matrices.

$$\hat{\Theta} = \arg\min_{\Theta} \sum_{k=0}^{N-1} (Y_k - H\Theta)^2 \tag{7.1}$$

$$\hat{\Theta} = (H^T H) \backslash (H^T Y) \tag{7.2}$$

This leaves us with the following estimations for all parameters.

$\hat{a}_{20a}$	$\hat{a}_{20b}$	$\hat{a}_{20c}$
$\hat{a}_1 = -0.8997$	$\hat{a}_1 = -0.8921$	$\hat{a}_1 = -1$
$\hat{a}_2 = -0.3599$	$\hat{a}_2 = 0.3302$	$\hat{a}_2 = 0.6$
$\hat{b}_0 = 0.0012$	$\hat{b}_0 = -6.7602 * 10^{-4}$	$\hat{a}_3 = -0.3$
	$\hat{b}1 = 0.9922$	$\hat{b}_1 = 0.9992$

7.2. b

We execute the predicted values from the models and compare the results with the validation data. We then use root mean square to get an error of the model.

$\hat{a}_{20a}$	$\hat{a}_{20b}$	$\hat{a}_{20c}$
$RMSE_{pred} = 1.0284$	$RMSE_{pred} = 2.0735$	$RMSE_{pred} = 1.0284$
$RMSE_{sim} = 1.4547$	$RMSE_{sim} = 2.0881$	$RMSE_{csim} = 2.1014$

7.3. c

$\hat{a}_{20a}$	$\hat{a}_{20b}$	$\hat{a}_{20c}$
$Cov = 2.1168$	$Cov = 4.3607$	$Cov = 4.4160$

The first model has the lowest errors, in both simulation and prediction. It also has the lowest covariance by far. This is therefore considered the superior model.