Parameter and State Estimation - Assignment 5

G Prashant (BS17B011)

January 2021

Question 1

Part (a)

Let the time varying model be written as,

$$y[k] = x_{\gamma_k}(k)\theta_{\gamma_k}$$

where $\gamma_k \in \{1, 2\}$, i.e., it represents which model is active at time instant k. Here, θ_1 represents the parameter vector for Model 1 and θ_2 represents the parameter vector for Model 2. Likewise, $x_1(k)$ represents the regressor vector for model 1 and $x_2(k)$ represents the regressor vector for model 2, at time instant k. For example, in the case of Model 1,

$$\theta_1 = \begin{bmatrix} a_1^{(1)} \\ b_1 \\ b_2 \end{bmatrix}, \quad x_1(k) = \begin{bmatrix} y[k-1] & u[k-1] & u[k-2] \end{bmatrix}$$

The problem can be formulated as that of a recursive least squares (RLS) by first initializing the parameters and covariances P for both models. Switch identification can be done as follows,

$$\gamma_k = \arg\min_{j=1,2}(y[k] - x_j(k)\theta_j[k-1])$$

The prediction error corresponding to model that gave minimum residual is

$$\epsilon[k] = y[k] - \hat{y}_{\gamma_k}[k] = y[k] - x_{\gamma_k}(k)\theta_{\gamma_k}[k]$$

The Parameter Covariance matrix is updated as follows

$$P_{\gamma_k}[k] = \frac{1}{\lambda}(P_{\gamma_k}[k-1] - \frac{P_{\gamma_k}[k-1]x_{\gamma_k}(k)^T x_{\gamma_k}(k)P_{\gamma_k}[k-1]}{\lambda + x_{\gamma_k}(k)P_{\gamma_k}[k-1]x_{\gamma_k}(k)^T})$$

The gain vector is updated as follows

$$K_{\gamma_k}[k] = \frac{P_{\gamma_k}[k-1]x_{\gamma_k}(k)^T}{\lambda + x_{\gamma_k}(k)P_{\gamma_k}[k-1]x_{\gamma_k}(k)^T}$$

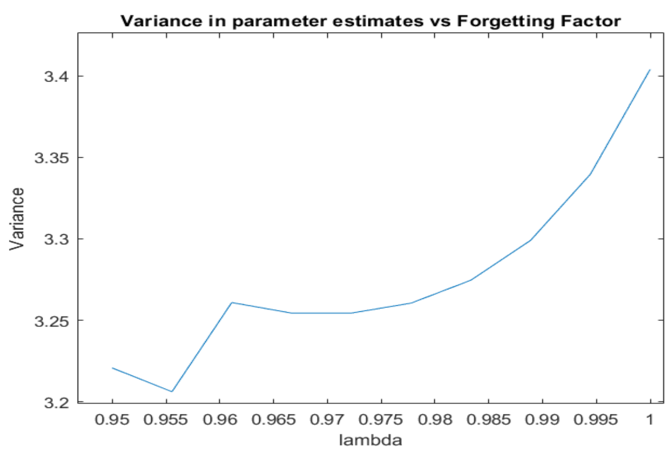
Here, λ is the forgetting factor. Finally, parameters corresponding to the best model are updated

$$\hat{\theta}_{\gamma_k}[k] = \hat{\theta}_{\gamma_k}[k-1] + K_{\gamma_k}[k]\epsilon[k]$$

In this question, one of the parameters b_2 is shared across Model 1 and Model 2. To overcome the constraint that this parameter should be same in both models, one approach would be to use the same estimate of b_2 as produced using Model 1 in Model 2. This way the, the parameters we estimate for Model 2 becomes $a_1^{(2)},\ a_2,\ b_3$, and the output variable becomes $y[k] - b_2u[k-2]$. For the subsequent parts in this question, this approach was followed.

Part (b)

Using the `recursiveLS` routine in MATLAB, the above formulation was executed and the forgetting factor λ was optimized by choosing one of 10 values that lie between 0.95 and 1 (inclusive), which gives the minimum residual squared error. The plot below shows variance of parameter estimates against increasing λ



It can be noticed that $\lambda = 0.9556$ gives the minimum variance. The parameter estimates are as follows,

$$a_1^{(1)} = 0.7869$$

$$a_1^{(2)} = 1.3246$$

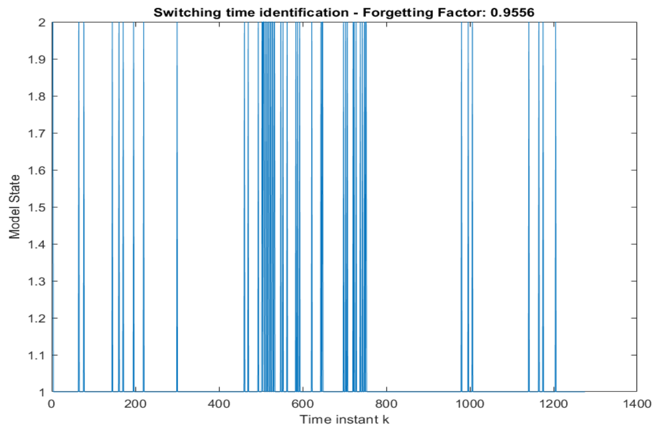
$$a_2 = 0.8496$$

$$b_1 = 1.9120$$

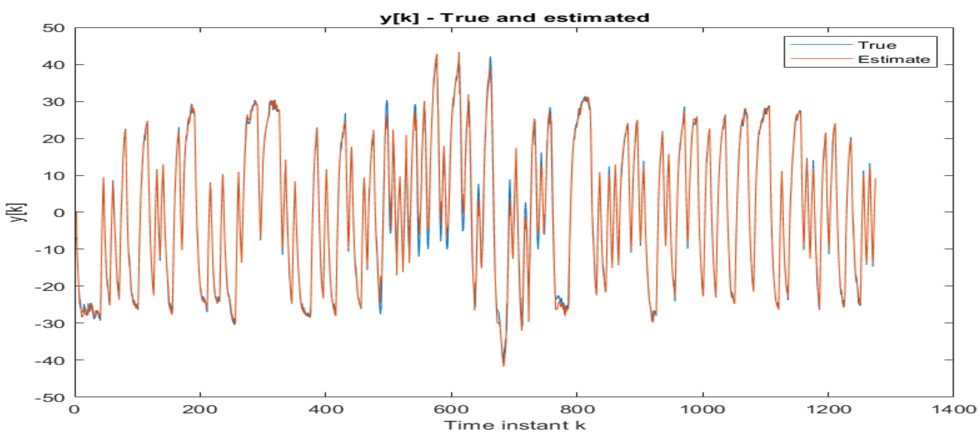
$$b_2 = 0.9770$$

$$b_3 = 1.1520$$

The variance of error for these estimates is 3.2062.
The plot below shows the switching time identification.

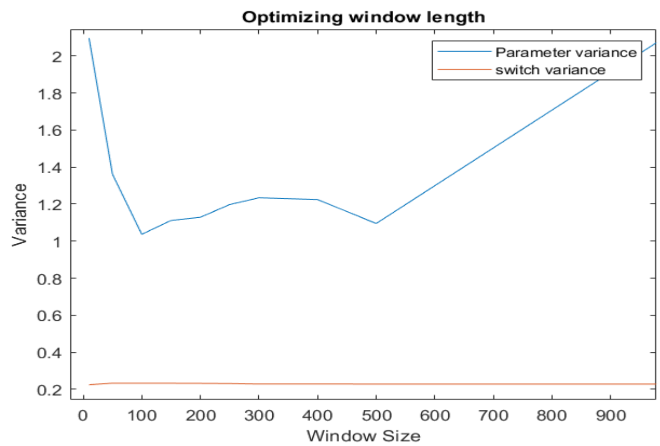


The plot below shows the the true and predicted values of $y[k]$.

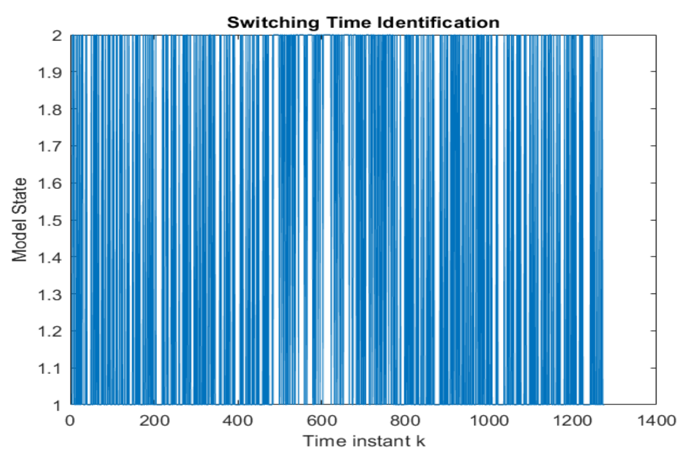


Part (c)

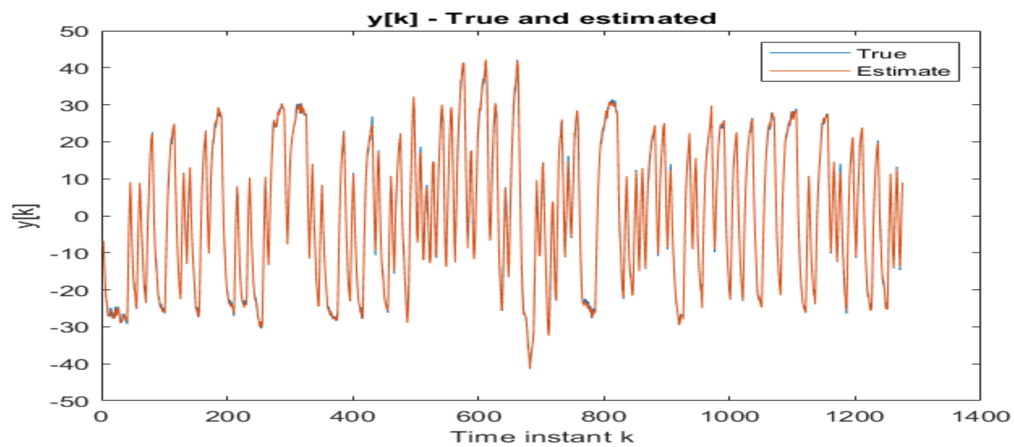
Similar to the previous question, the window length was optimized for a trade-off between variance of parameter estimates and identification of switching times. The finite history option in RLS was chosen for this purpose.
The plot below shows the variation in variance of parameter estimates and switching identification times upon increasing window size - from 10 to 1000.



Therefore, it can be noticed that the variance in identification of switching times is nearly undisturbed with varying window size. However, the variance of parameter estimates takes a minimum value of 1.0370 when the window length is 100. The plot below shows the switching time identification.



Evidently, this approach predicts that the model switches much more frequently than in the forgetting factor approach. The plot below shows the the true and predicted values of $y[k]$.



Part (d)

From the above observations, we can say that the forgetting factor approach results in higher variance of parameter estimates but lower variance in identification of switching times. On the other hand, in the case of moving window approach the variance in parameter estimates is much lower but that of identification of switching times is higher. This suggests that the moving window model is much sensitive to changes in model state, while producing more accurate estimates of output. Hence, the moving window approach is more suitable for the given data than the forgetting factor approach.

Question 2

The state space model description is given as

$$A = \begin{bmatrix} 0.9 & 0 & 0 \\ 1 & 1.2 & -0.5916 \\ 0 & 0.5916 & 0 \end{bmatrix},\ B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ C = \begin{bmatrix} 2 & 0.8 & -0.6761 \end{bmatrix},\ D = 0$$

Part (a)

For the given state space model to be a minimal realization, both the observability matrix and the controllability matrix should be full rank.

The observability matrix is given by

$$\mathcal{O}_3 = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 2 & 0.8 & -0.6761 \\ 2.6 & 0.56 & -0.4733 \\ 2.9 & 0.392 & -0.3313 \end{bmatrix}$$

The echlon form of the observability matrix is

$$\text{echelon}(\mathcal{O}_3) \approx \begin{bmatrix} 2 & 0.8 & -0.6761 \\ 0 & -0.369 & -0.312 \\ 0 & 0 & 0 \end{bmatrix}$$

The rank of $\mathcal{O}_3$ is 2.

Similarly, the controllability matrix is given by

$$\mathcal{C}_3 = \begin{bmatrix} B & BA & BA^2 \end{bmatrix} = \begin{bmatrix} 1 & 0.9 & 0.81 \\ 0 & 1 & 2.1 \\ 0 & 0 & 0.5916 \end{bmatrix}$$

The rank of $\mathcal{C}_3$ is 3.

From the above analyses, the system is controllable but not observable, suggesting that it is not a minimal realization.

Using the `obsvf` command in MATLAB, the cannonical observation decomposition of the system is given below:

$$\bar{A} = \begin{bmatrix} 0.5 & -0.7487 & 1.1208 \\ 0 & 0.4292 & -0.5172 \\ 0 & 0.2466 & 1.1709 \end{bmatrix}, \ \bar{B} = \begin{bmatrix} 0 \\ 0.4639 \\ 0.8859 \end{bmatrix}, \ \bar{C} = \begin{bmatrix} 0 & 0 & 2.2577 \end{bmatrix},$$

The observable components are therefore,

$$A_o = \begin{bmatrix} 0.4292 & -0.5172 \\ 0.2466 & 1.1709 \end{bmatrix}, \ B_o = \begin{bmatrix} 0.4639 \\ 0.8859 \end{bmatrix}, \ C_o = \begin{bmatrix} 0 & 2.2577 \end{bmatrix}, \ D_o = 0$$

Therefore, the above system is a minimal realization of the highest order, with two states.

Part (b)

In order to place the eigenvalues the above minimal realization in the interval [0.05,3], we place the eigenvalues as 0.1 and 0.2, and determine the gain K using `place` function in MATLAB.

$$\lambda(A_o - KC_o) = (0.1, 0.2)$$

The gain is,

$$K = \begin{bmatrix} -0.0936 \\ 0.5758 \end{bmatrix}$$

The observer of the system is therefore,

$$\hat{x}[k+1] = A_o\hat{x}[k] + B_o u[k] + K(y[k] - C_o\hat{x}[k])$$

$$\hat{y}[k] = C_o\hat{x}[k] + D_o u[k]$$

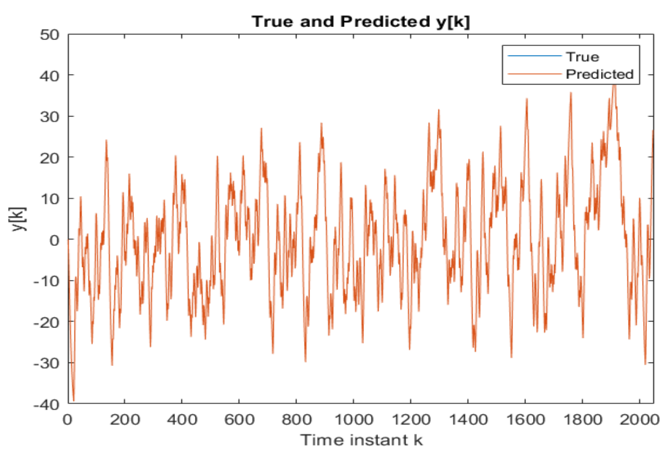
or

$$\hat{x}[k+1] = \begin{bmatrix} 0.4292 & -0.5172 \\ 0.2466 & 1.1709 \end{bmatrix} \hat{x}[k] + \begin{bmatrix} 0.4639 \\ 0.8859 \end{bmatrix} u[k] + K(y[k] - \begin{bmatrix} 0 & 2.2577 \end{bmatrix} \hat{x}[k])$$

$$\hat{y}[k] = \begin{bmatrix} 0 & 2.2577 \end{bmatrix} \hat{x}[k]$$

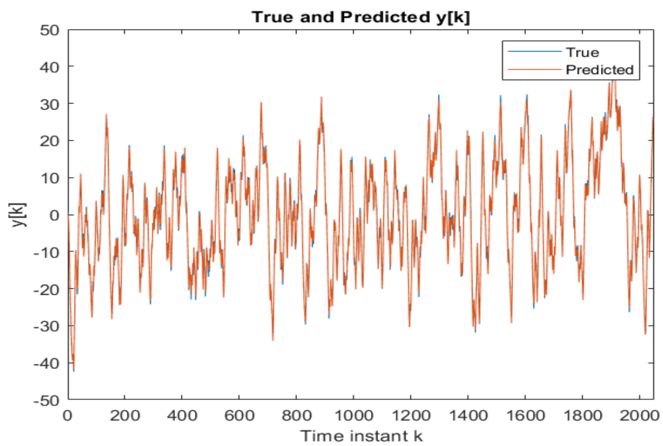
Part (c)

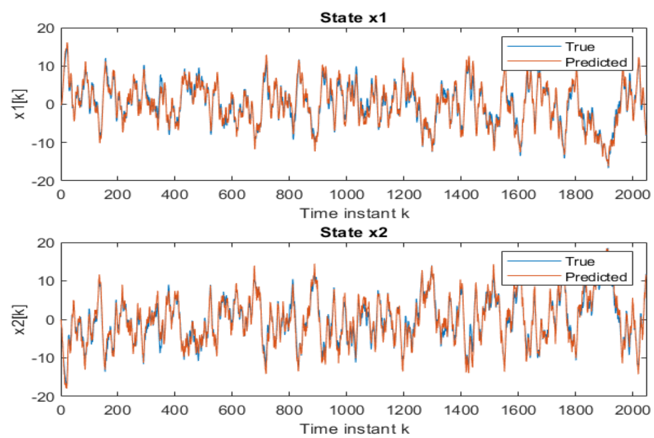
Input-Output data was generated using the original non-observable model given in the question. Following this, the observer of the minimal realization model was designed with the gain vector (using estim command) was implemented to verify prediction performance. It was observed that almost nearly all the values were predicted accurately, and the variance of measurement estimates was of the order of 10^{-9}. The plot below shows that the predicted and true values overlap.



Part (d)

The above procedure was repeated with process and measurement noise, with variances 0.1 and 1, respectively. The corresponding fit is shown below, and the variance of state estimates was found to be 3.2634 and variance of output estimates was found to be 1.4452.





When any one of the eigenvalue is greater than 1 in magnitude, the system becomes unstable and the states do not get estimated. When both eigenvalues are less than 1 in magnitude, the system is stable and the states are estimated appropriately. In the latter scenario, it was determined through simulation that the variance of the state estimates increased when the magnitude of the eigenvalues increased.

Question 3

Part (a)

From the given diagram, we can write the rate of change of h_2 as follows:

$$\frac{dh_2}{dt} = (F_{o1} - F_{o2})/A_2 \tag{1}$$

Similarly, the rate of change of h_1 is given as follows:

$$\frac{dh_1}{dt} = (F_i - F_{o1})/A_1 \tag{2}$$

Also, as the flow rates are linearly proportional to the heights,

$$F_{o1} = C_{v1}h_1 \tag{3}$$

$$F_{o2} = C_{v2}h_2 \tag{4}$$

From equations (3) and (4), we get

$$\frac{dh_1}{dt} = \frac{F_i}{A_1} - \frac{C_{v1}h_1}{A_1} \tag{5}$$

and

$$\frac{dh_2}{dt} = \frac{C_{v1}h_1}{A_2} - \frac{C_{v2}h_2}{A_2} \tag{6}$$

The continuous time state-space model can be formulated from the above expressions as follows

$$\underbrace{\begin{bmatrix} \frac{dh_1}{dt} \\ \frac{dh_2}{dt} \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} -\frac{C_{v1}}{A_1} & 0 \\ \frac{C_{v1}}{A_1} & -\frac{C_{v2}}{A_2} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} \frac{1}{A_1} \\ 0 \end{bmatrix}}_{B} \underbrace{F_i}_{u} \tag{7}$$

and the output equation is as follows (assuming F_{o2} is the output)

$$y = F_{o2} = \underbrace{\begin{bmatrix} 0 & C_{v2} \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}}_{x} \tag{8}$$

Analogously the state space model is given by

$$\dot{x} = Ax + Bu \tag{9}$$
$$y = Cx \tag{10}$$

Part (b)

After descretizing the model with `c2d` command, the following model was obtained

$$x[k+1] = \begin{bmatrix} 0.9355 & 0 \\ 0.06672 & 1.069 \end{bmatrix} x[k] + \begin{bmatrix} 0.08062 \\ 0.002779 \end{bmatrix} u[k] \tag{11}$$
$$y[k] = \begin{bmatrix} 0 & 0.8 \end{bmatrix} x[k] \tag{12}$$

Part (c)

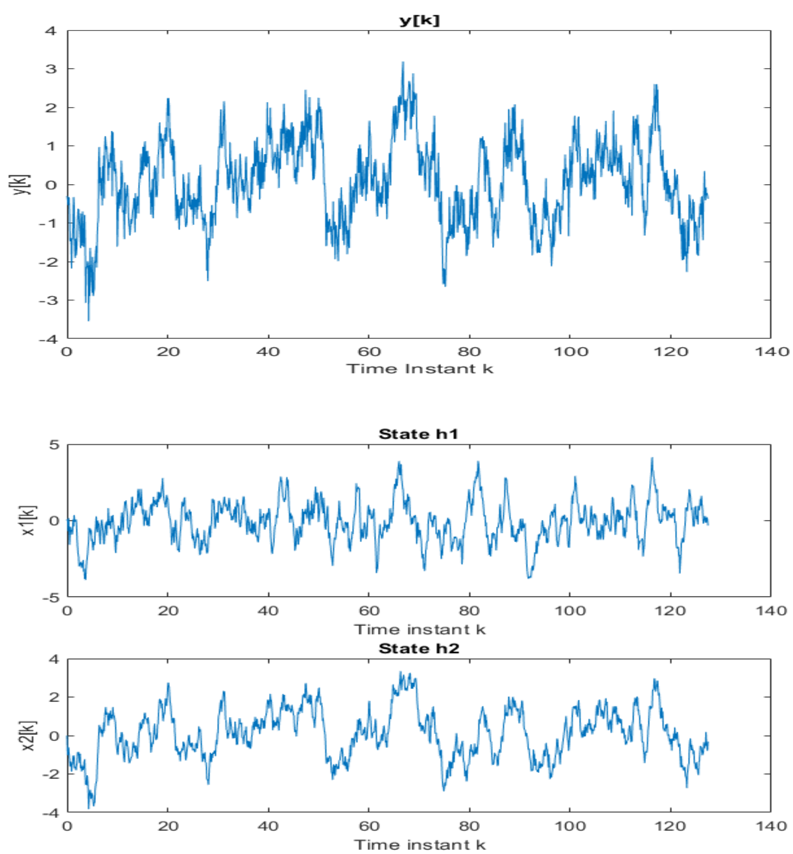
In equation (8), we had already assumed that output F_{o2} is only sensed. In this case, the observability matrix (from continuous time state-space model) is given as follows:

$$\mathcal{O}_2 = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 0 & 0.8 \\ 0.5333 & 0.5333 \end{bmatrix}$$

As noticable, the rank of $\mathcal{O}_2$ is 2, which means the system is observable. Hence, both the states $h_1,\ h_2$ can be uniquely determined in noise-free conditions.

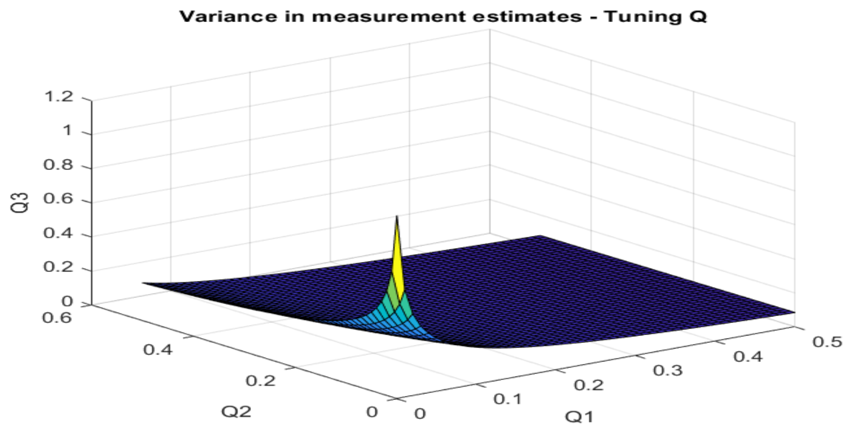
Part (d)

As given in the question, input-output data was generated from the discrete state-space model, as per the directions given in the question - process noise variance Q and measurement noise variance R. The plot of the simulated data is shown below:

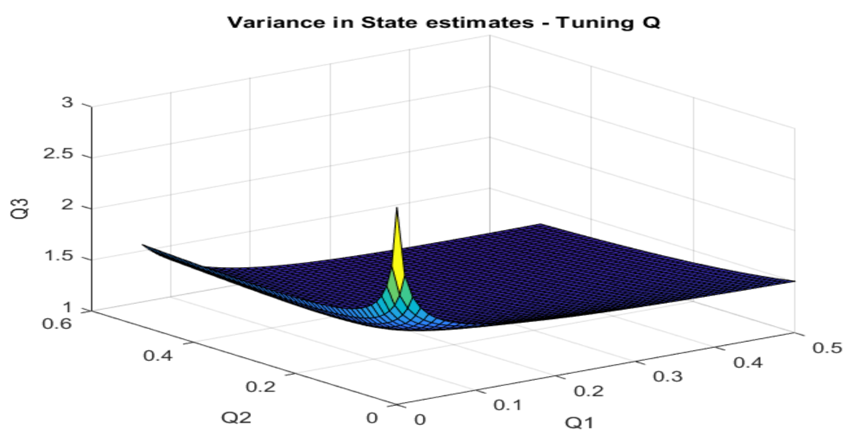


Part (e)

Kalman filter was implemented for estimating the states $h_1,\ h_2$ using the data generated in (3d). The diagonal elements of the $\boldsymbol{Q}$ matrix (Q_1, Q_2) were tuned to lie between 0 and 0.5. The plot below shows the variance of the output estimate with varying $Q_1,\ Q_2$ values.



From this plot, it is clearly evident that increasing both Q_1 and Q_2 decreases the variance of the prediction error, i.e., error in output estimate. However, it is recommended that the variance in the state estimates are also checked. The plot below shows the same.



From the above plot of variance of errors of state estimates, it is noticable that increasing Q_1 continuously results in decreases in variance, while in the case of Q_2, the variance of the state estimates reach a minimum value of 1.3. The values of $Q_1,\ Q_2$ that gave produced the minimum value are

$$Q_1 = 0.5$$

$$Q_2 = 0.42$$

Therefore, the optimal $\boldsymbol{Q}$ is

$$\boldsymbol{Q} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.42 \end{bmatrix}$$

which is very different from the true $\boldsymbol{Q}$ matrix, which is $\begin{bmatrix} 0.2 & 0 \\ 0 & 0.1 \end{bmatrix}$

This indicates that tuned $\boldsymbol{Q}$ does not match with the true matrix. Moreover, Q_2 has more influence on the variance of the state estimates that Q_1. This can be attributed to the fact that the measured output is only dependent on h_2.

Question 4

In this question, we are given that the flow rates of each tank are proportional to the square roots of the heights of water present in the corresponding tank. This means,

$$F_{o1} = C_{v1}\sqrt{h_1} \tag{13}$$

$$F_{o2} = C_{v2}\sqrt{h_2} \tag{14}$$

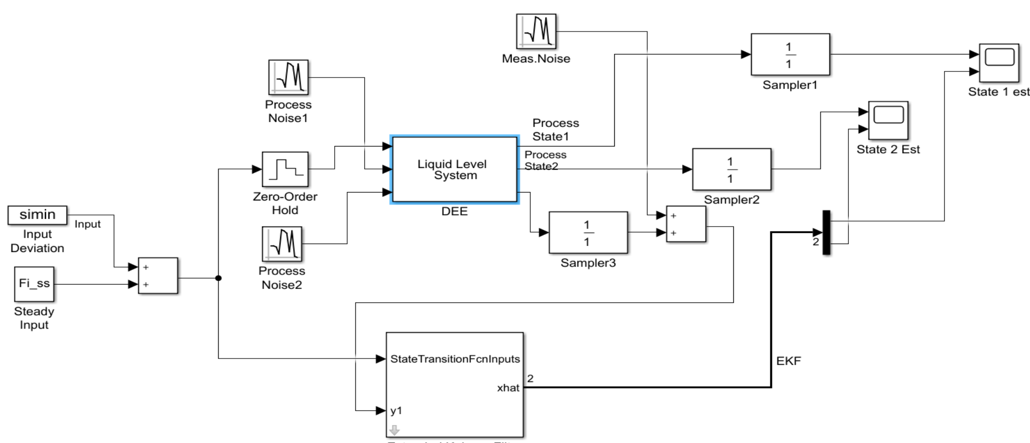
The state equation therefore becomes

$$\frac{dh_1}{dt} = \frac{F_i}{A_1} - \frac{C_{v1}\sqrt{h_1}}{A_1} \tag{15}$$

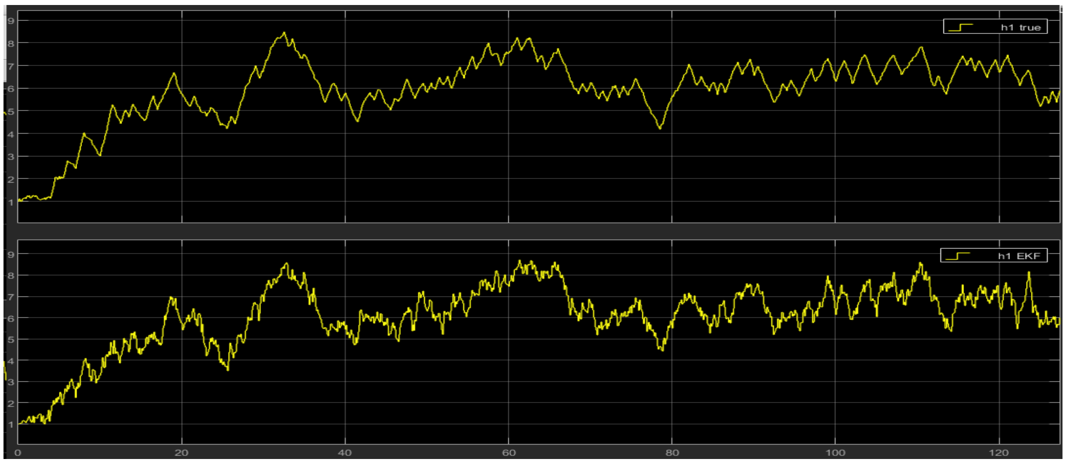
and

$$\frac{dh_2}{dt} = \frac{C_{v1}\sqrt{h_1}}{A_2} - \frac{C_{v2}\sqrt{h_2}}{A_2} \tag{16}$$

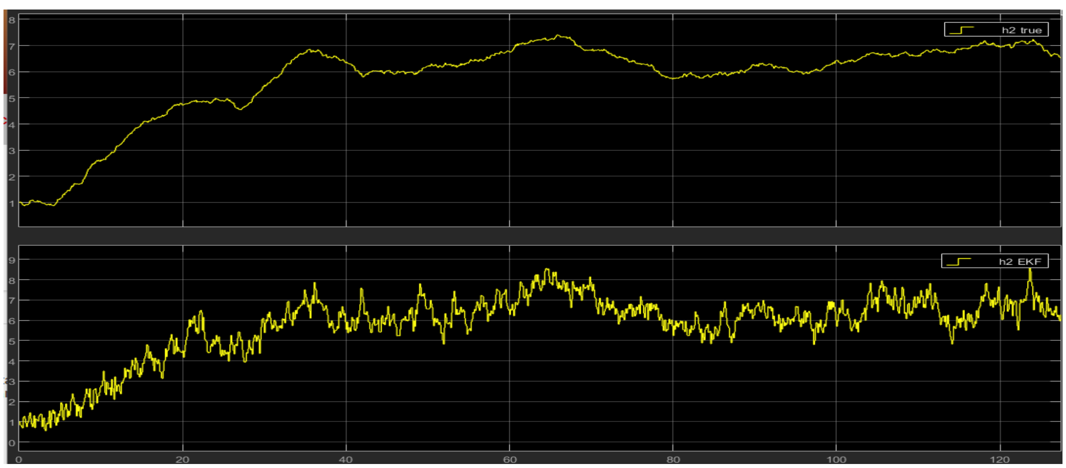
This is clearly a nonlinear system, and an Extended Kalman Filter (EKF) is more suitable for handling these systems. EKF was implemented on SIMULINK for this state system. The block diagram is given below:



Both the state values were initialized to 1 during simulation and in EKF. The parameter covariance P was initialized to $10^{-4}I$. The plot showing true and EKF estimates of State h_1 is shown below:



Likewise, the plot showing true and EKF estimates of State h_2 is shown below:



The SIMULINK Model is attached with the submission with the name EKF.slx

Question 5

We are given that we have $N+1$ noisy observations of X.

$$y[k] = x[k] + v[k]$$

The state space equation is given by

$$x[k+1] = \mu + w[k]$$

where,

$$w[k] \sim \mathcal{N}(0, \sigma_X^2)$$

The state space matrices are therefore given by

$$A = 0,\ B = 1,\ C = 1,\ D = 0$$

The measurement update equation is

$$\hat{x}[k|k] = \hat{x}[k|k-1] + K_{f,k} e[k]$$

$$K_{f,k} = \frac{P[k|k-1]}{\sigma_v^2 + P[k|k-1]}$$

$$P[k|k] = \frac{\sigma_v^2 P[k|k-1]}{\sigma_v^2 + P[k|k-1]}$$

$$\hat{x}[k+1|k] = \mu$$

$$P[k+1|k] = \sigma_X^2$$