TP Kalman - Master SIVOS 2024, Université de Rennes
Kalman filter Implementation

Quinteros Santiago

I. INTRODUCTION

In this article we will simulate a Kalman filter with invariant matrices F, G and H and known noise variances Q and R. In the first section a Kalman filtering of a 4D-state 2D-observation system will be simulated in octave changing the initial conditions. In the second one, the case of a non-zero-mean Kalman model will be analyzed.

II. KALMAN FILTERING OF A 4D-STATE 2D-OBSERVATION SYSTEM

System:

$$\begin{cases} \boldsymbol{X}_{k+1} = F\boldsymbol{X}_k + G\boldsymbol{U}_k \\ \boldsymbol{Y}_k = H\boldsymbol{X}_k + \boldsymbol{B}_k \end{cases} \tag{1}$$

Assumptions on the up to second-order signal statistics are:

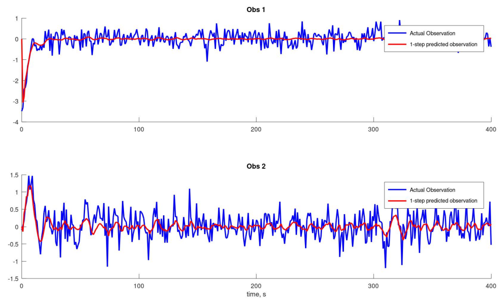
$$\begin{gathered} E\left[\boldsymbol{X}_0\right] = \overline{\boldsymbol{x}}_0, \quad E\left[\boldsymbol{B}_k\right] = \mathbf{0}, \quad E\left[\boldsymbol{U}_k\right] = \mathbf{0} \\ E\left(\left[\begin{array}{c} \boldsymbol{B}_k \\ \hline \boldsymbol{X}_0 \\ \hline \boldsymbol{U}_k \end{array}\right]\left[\boldsymbol{B}_l^{\mathrm{t}}\left|\boldsymbol{X}_0^{\mathrm{t}}\right|\boldsymbol{U}_l^{\mathrm{t}}\right]\right) = \left[\begin{array}{c|c|c} \mathbf{R}_k\delta_{kl} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{P}_0 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{Q}_k\delta_{kl} \end{array}\right] \end{gathered} \tag{2}$$

Therefore the Kalman filter implementation is as follows:

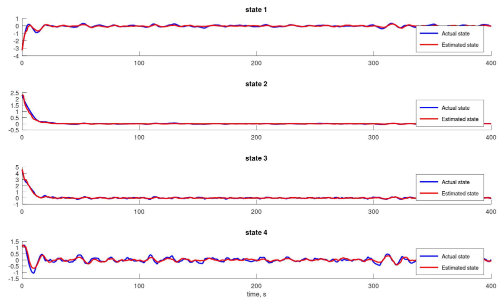
$$\begin{gathered} \hat{\boldsymbol{X}}_{k|\mathcal{Y}_k} = F\hat{\boldsymbol{X}}_{k-1|k-1} + \boldsymbol{\Gamma}_k(\boldsymbol{Y}_k - \hat{\boldsymbol{Y}}_{k|k-1}) \\ \hat{\boldsymbol{Y}}_{k|k-1} = H\hat{\boldsymbol{X}}_{k|k-1} \\ \hat{\boldsymbol{X}}_{k|k-1} = F\hat{\boldsymbol{X}}_{k-1|k-1} \\ \\ \boldsymbol{\Gamma}_k = \mathbf{C}_{\hat{\boldsymbol{X}}_k\alpha_k}\Sigma_k^{-1} = \mathbf{K}_{k|k-1}H^T\left(\mathbf{R} + H\mathbf{K}_{k|k-1}H^T\right)^{-1} \\ \mathbf{K}_{k+1|k} = F\mathbf{K}_{k|k}F^T + G\mathbf{Q}G^T \\ \mathbf{K}_{k|k} = \left(\mathbf{I} - \boldsymbol{\Gamma}_kH\right)\mathbf{K}_{k|k-1} \end{gathered} \tag{3}$$

x0= 4*randn(4,1)

A. *Initial results with x1=x0= 4*randn(4,1), known initial state*



(a) Observations and predictions.



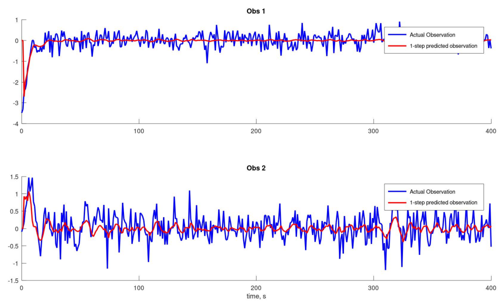
(b) Real and estimated states evolution

Fig. 1: Initial results with x1=x0, known initial state

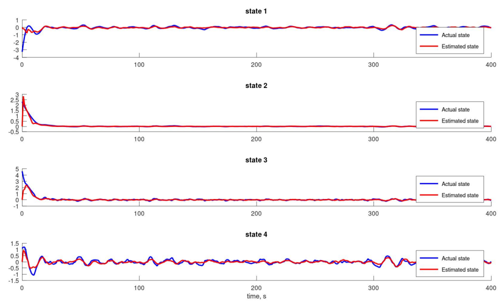
The good quality of the state estimation can be appreciated and that the estimated observation is smoother than the real one to avoid an innovation value equal to zero. We use the difference between the estimated observation and the actual one to calculate the innovation.

B. *Question 1: Change the initial states x1=x0, what are the consequences?*

I set x1= randn(4,1) as the initial state estimation of the kalman filter.



(a) Observations and predictions.



(b) Real and estimated states evolution

Fig. 2: Results with x1≠x0, x1= randn(4,1)

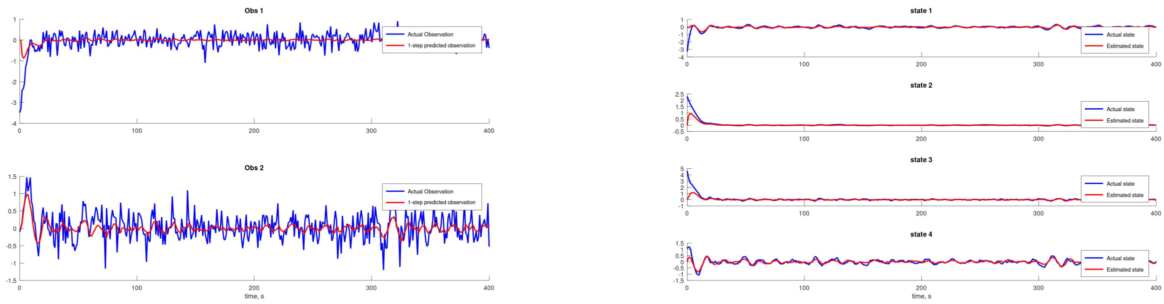
Even if at the beginning the state estimations are not close to the actual value, they rapidly converge to it. Thus showing the power of the Kalman filter.

C. *Question 2: Change the initial state covariance and observe the convergence rate as a function of these variances and explain the phenomenon*

If the initial state covariance is higher, there is a high uncertainty associated with the initial state of the system. This implies an **increased variability in predictions**: During the initial stages of the prediction process, the large initial covariance will allow for greater variability in the Kalman filter predictions. This means that predictions of the system state could be less accurate and more dispersed, reflecting the high initial uncertainty. We also have **more iterations to converge**: Due to the high initial uncertainty, the Kalman filter may require more iterations to converge to an accurate estimate of the system state.

The updating and prediction process may require more time and observations to reduce the initial uncertainty and improve the accuracy of the state estimation.

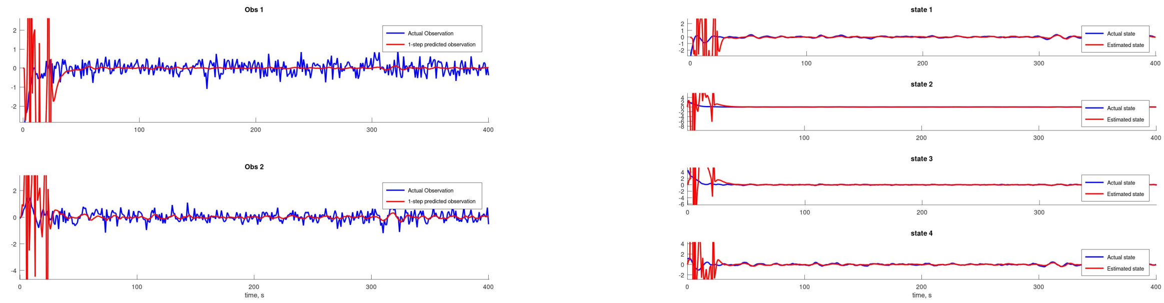
This effects can be appreciated in figures 3 and 4 where P1 are 0.04*eye(4) and 40*eye(4) respectively.



(a) Observations and predictions.

(b) Real and estimated states evolution

Fig. 3: Results with x1≠x0, x1= randn(4,1), P1=0.04*eye(4)



(a) Observations and predictions.

(b) Real and estimated states evolution

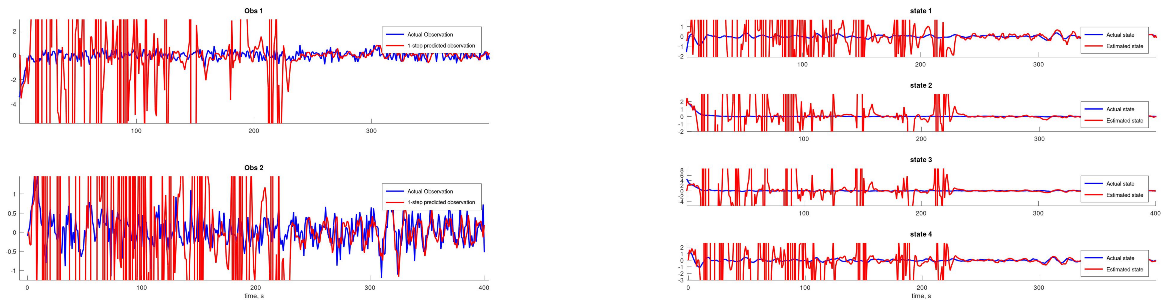
Fig. 4: Results with x1≠x0, x1= randn(4,1), P1=40*eye(4)

D. Question 3: Introduce uncertainties to the process parameters such that F1=F,G1=G,H1=H(one for each test)are not known exactly to the Kalman filter. Describe the tolerances of the filter wrt these uncertainties.

This section will show the importance of a good model and the impacts of not having a good one.

F1:

F1=F+0.25*randn(size(F)). When F1 is uncertain to the filter, then the predictions of both observations and states are worse, mostly in the beginning.



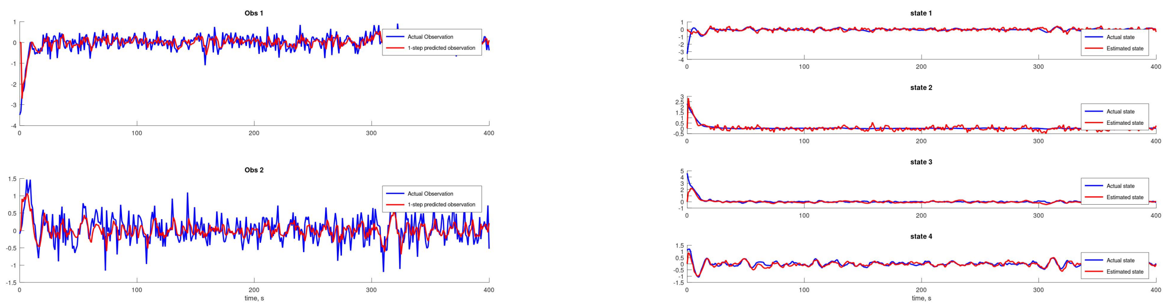
(a) Observations and predictions.

(b) Real and estimated states evolution

Fig. 5: Results with x1≠x0, x1= randn(4,1), F1=F+0.25*randn(size(F))

G1:

G1=G+0.25*randn(size(F)). When G1 is uncertain to the filter, no big impacts are seen, this makes sense as it is just affecting the impact of the process noise. The observation estimation is less smooth tha before.



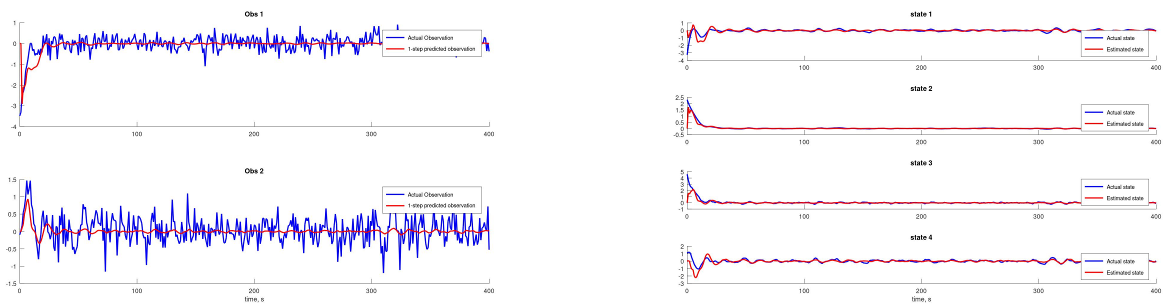
(a) Observations and predictions.

(b) Real and estimated states evolution

Fig. 6: Results with x1≠x0, x1= randn(4,1), G1=G+0.25*randn(size(G))

H1:

H1=H+0.5*randn(size(H)). When H1 is uncertain to the filter, predicted x0 are slightly worse.



(a) Observations and predictions.

(b) Real and estimated states evolution

Fig. 7: Results with x1≠x0, x1= randn(4,1), H1=H+0.5*randn(size(H))

III. Case of non-zero-mean Kalman model

A. *Question 1*

We aim to extend the previous Kalman model to the more general non-zero-mean stochastic case ($E[U'_k] \neq$, $E[B'_k] \neq 0$),all other assumptions being identical.

We define a $W_k = U_k + E[U'_k]$ and $V_k = B_k + E[B'_k]$

System:

$$\begin{cases} \boldsymbol{X}_{k+1} = F_k \boldsymbol{X}_k + G_k \boldsymbol{W}_k \\ \boldsymbol{Y}_k = H_k \boldsymbol{X}_k + \boldsymbol{V}_k \end{cases} \tag{4}$$

Consequently

$$\begin{cases} \boldsymbol{X}_{k+1} = F_k \boldsymbol{X}_k + \boldsymbol{C}_k + G_k \boldsymbol{U}_k \\ \boldsymbol{Y}_k = H_k \boldsymbol{X}_k + \boldsymbol{D}_k + \boldsymbol{B}_k \end{cases} \tag{5}$$

Assumptions on the up to first-order signal statistics are:

$$E\left[\boldsymbol{X}_0\right] = \overline{\boldsymbol{x}}_0, \quad E\left[\boldsymbol{B}_k\right] = \boldsymbol{0}, \quad E\left[\boldsymbol{U}_k\right] = \boldsymbol{0}, \quad E\left[\boldsymbol{C}_k\right] = G_k E[U'_k], \quad E\left[\boldsymbol{D}_k\right] = E_k[B'_k] \tag{6}$$

B. *Question 2: Explain that $E[X_k], E[Y_k]$ are known and calculable without any observations*

$E[X_k] = F_{k-1}E[X_{k-1}] + C_{k-1}$

$E[X_{k-1}] = F_{k-2}E[X_{k-2}] + C_{k-2}$

...

$E[X_0] = \overline{x_0}$

$So, iteratively:$

$E[X_k] = (F_{k-1}...F_0)\overline{x_0} + C_{k-1} + F_{k-1}C_{k-2} + F_{k-1}F_{k-2}C_{k-3} + ...$

And

$E[Y_k] = H_k E[X_k] + D_k$

where $E[X_k]$ is the one calculated above.

C. *Question 3: Show that the innovation process $\alpha_k = \boldsymbol{Y}_k - \hat{\boldsymbol{Y}}_{k|k-1}$ remain centered*

We remember the expression of X_k:

$$\boldsymbol{X}_k = F_{k-1}\boldsymbol{X}_{k-1} + \boldsymbol{C}_{k-1} + G_{k-1}\boldsymbol{U}_{k-1}$$

The general expression of a LMMSE estimator:

$$\hat{\theta}_{LMMSE} = E[\theta] + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(x - E[x])$$

In particular:

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

Remembering:

$$\hat{\boldsymbol{X}}_{i|n} \text{ : estimateur optimal linéaire de } \boldsymbol{X}_i \text{ sachant } \mathcal{Y}_n \tag{7}$$

So, taking:

$$\begin{aligned} \hat{\theta} &= \hat{\boldsymbol{X}}_{k|k-1} \\ \theta &= X_k \\ x &= \mathcal{Y}_{k-1} \end{aligned} \tag{8}$$

We get :

$$E[\hat{\boldsymbol{X}}_{k|k-1}] = E[X_k]$$

Equivalently for:

$$\begin{aligned} \hat{\theta} &= \hat{\boldsymbol{X}}_{k|k} \\ \theta &= X_k \\ x &= \mathcal{Y}_k \end{aligned} \tag{9}$$

We get :

$$E[\hat{\boldsymbol{X}}_{k|k-1}] = E[\hat{\boldsymbol{X}}_{k|k}] = E[X_k]$$

For the case of Y, we remember the formula:

$$\boldsymbol{Y}_k = H_k \boldsymbol{X}_k + \boldsymbol{D}_k + \boldsymbol{B}_k \tag{10}$$

Taking:

$$\begin{aligned} \hat{\theta} &= \hat{\boldsymbol{Y}}_{k|k-1} \\ \theta &= Y_k \\ x &= \mathcal{Y}_{k-1} \end{aligned} \tag{11}$$

We get :

$$E[\hat{\boldsymbol{Y}}_{k|k-1}] = E[Y_k]$$

Thus the innovation process $\alpha_k = \boldsymbol{Y}_k - \hat{\boldsymbol{Y}}_{k|k-1}$ remains centered.

D. Question 4: show that $\mathcal{Y}_k = \{\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_k\}$ can be linearly expressed with $\{\alpha_1, \ldots, \alpha_k\}$ knowing $E[\boldsymbol{X}_k], E[\boldsymbol{Y}_k]$; as a consequence $\hat{\boldsymbol{X}}_{k|\mathcal{Y}_k} = \hat{\boldsymbol{X}}_{k|\alpha_1,\ldots,\alpha_k}$ and is noted as $\hat{\boldsymbol{X}}_{k|k}$ for simplicity;

1) $\{\alpha_1, ..., \alpha_k\}$ is a LC of $\mathcal{Y}_k = \{Y_1, ..., Y_k\}$:
By definition: $\alpha_k = \boldsymbol{Y}_k - \hat{\boldsymbol{Y}}_{k|k-1}$ and as $\hat{\boldsymbol{Y}}_{k|k-1}$= $\hat{\boldsymbol{Y}}_{k|\mathcal{Y}_{k-1}}$ is a linear combination of $\mathcal{Y}_{k-1}$ then $\{\alpha_1, ..., \alpha_k\}$ is a LC of $\mathcal{Y}_k = \{Y_1, ..., Y_k\}$.

2)$\mathcal{Y}_k = \{Y_1, ..., Y_k\}$ is a LC of $\{\alpha_1, ..., \alpha_k\}$:
By definition: $\boldsymbol{Y}_k = \alpha_k - \hat{\boldsymbol{Y}}_{k|k-1}$=$\alpha_k - \hat{\boldsymbol{Y}}_{k|\mathcal{Y}_{k-1}}$. As shown in 1, if $\mathcal{Y}_{k-1}$ is a LC of $\{\alpha_1, ..., \alpha_{k-1}\}$ then $\mathcal{Y}_k$ is a LC of $\{\alpha_1, ..., \alpha_k\}$. This is clearly true in the case of $\alpha_1 = Y_1 - \hat{Y}_{1|0}$. By taking the expected value, and knowing the estimator is unbiased, it is clear that $\alpha_1 = Y_1 - E[Y_1]$, so $Y_1 = \alpha_1 - E[Y_1]$ thus $\mathcal{Y}_k = \{Y_1, ..., Y_k\}$ is a LC of $\{\alpha_1, ..., \alpha_k\}$

E. Question 5: $\hat{\boldsymbol{Y}}_{k|k-1} = \mathrm{H}_k \hat{\boldsymbol{X}}_{k|k-1} + \mathrm{D}_k$ and $\hat{\boldsymbol{X}}_{k+1|k} = \mathrm{F}_k \hat{\boldsymbol{X}}_{k|k} + \mathrm{C}_k$

$$\hat{X}_{k+1|k} = E[X_{k+1} \mid \mathcal{Y}_k] = E[F_k X_k + C_k + G_k U_k \mid \mathcal{Y}_k] = F_k E[X_k \mid \mathcal{Y}_k] + C_k = F_k \hat{X}_{k|k} + C_k$$

$$\hat{Y}_{k|k-1} = E[Y_k \mid \mathcal{Y}_{k-1}] = E[H_k X_k + D_k + B_k \mid \mathcal{Y}_{k-1}] = H_k E[X_k \mid \mathcal{Y}_{k-1}] + D_k = H_k \hat{X}_{k|k-1} + D_k \rightarrow$$

$$\hat{Y}_{k|k-1} = H_k F_{k-1} \hat{X}_{k-1|k-1} + D_k$$

F. Question 6: Recursive KF process

The objective is to find the expression of: $\hat{X}_{k|k}$
Using the result of the previous subsection we can affirm that:

$$\hat{X}_{k|k} = E[X_k \mid \mathcal{Y}_k] = E[X_k \mid \mathcal{Y}_{k-1}, \alpha_k] = E[X_k \mid \mathcal{Y}_{k-1}] + E[X_k \mid \alpha_k]$$

Notice that the following property is valid as α is centered: $E[\theta \mid x_1, x_2] = E[\theta \mid x_1] + E[\theta \mid x_2]$

Using the definition of X_n, the first part of the equation is:

$$E[X_k \mid \mathcal{Y}_{k-1}] = E[F_{k-1} X_{k-1} + C_{k-1} + G_{k-1} U_{k-1} \mid \mathcal{Y}_{k-1}] = F_{k-1} E[X_{k-1} \mid \mathcal{Y}_{k-1}] + C_{k-1} = F_{k-1} \hat{X}_{k-1|k-1} + C_{k-1}$$

Using the definition of LMMSE for α_k and as it is centered, the second part of the equation is:

$$E[X_k \mid \alpha_k] = C_{X_k \alpha_k} \Sigma_k^{-1} \alpha_k = \Gamma_k \alpha_k$$

Where:
$\alpha_k = Y_k - \hat{Y}_{k|k-1} = Y_k - E[H_k X_k + D_k \mid \mathcal{Y}_{k-1}] = Y_k - H_k \hat{X}_{k|k-1} - D_k = Y_k - H_k \left(F_{k-1} \hat{X}_{k-1|k-1} + C_{k/1} \right) - D_k \rightarrow$
$\alpha_k = Y_k - H_k F_{k-1} \hat{X}_{k-1|k-1} - H_k C_{k/1} - D_k$ Then:
$\hat{X}_{k|k} = F_{k-1} \hat{X}_{k-1|k-1} + C_{k-1} + \Gamma_k (Y_k - H_k F_{k-1} \hat{X}_{k-1|k-1} - H_k C_{k/1} - D_k)$

G. Question 7:Γ_k deduction

$$\Gamma_k = C_{X_k \alpha_k} \Sigma_k^{-1}$$

Using the definition of $\hat{Y}_{k|k-1}$ and that $\hat{X}_{k|k-1}$ is uncorrelated with $Y_k - \hat{Y}_{k|k-1}$ then

$C_{X_k \alpha_k} = \text{Cov}\left(X_k, Y_k - \hat{Y}_{k|k-1}\right) = \text{Cov}\left(X_k - \hat{X}_{k|k-1}, Y_k - \hat{Y}_{k|k-1}\right)$

$C_{X_k \alpha_k} = \text{Cov}(X_k - \hat{X}_{k|k-1}, H_k(X_k - \hat{X}_{k|k-1} + B_k)$ As $X_k - \hat{X}_{k|k-1}$ and $B_k$are not correlated then:

$C_{X_k \alpha_k} = var(X_k - \hat{X}_{k|k-1})H^T{}_k = K_{k|k-1}H^T{}_k$

$\Sigma_k = E\left[\alpha_k \alpha_k^T\right] = var(X_k - \hat{X}_{k|k-1}, H_k(X_k - \hat{X}_{k|k-1} + B_k)) = H_k var(X_k - \hat{X}_{k|k-1})H^T{}_k + var(B_k) \rightarrow$

$\Sigma_k = H_k K_{k|k-1} H^T{}_k + R_k$

Which means the Kalman Gain has the same shape as the zero-mean model:

$\Gamma_k = C_{\hat{X}_k \alpha_k} \Sigma_k^{-1} = K_{k|k-1} H_k^T \left(R_k + H_k K_{k|k-1} H_k^T\right)^{-1}$

Analysing the expressions of $K_{k|k-1}$ and $K_{k|k}$ we can conclude nothing is to be changed to generate the Kalman gains:

$K_{k|k-1} = var(X_k - \hat{X}_{k|k-1}) = E[(F_{k-1}(X_k - \hat{X}_{k|k-1}) + G_{k-1}U_{k-1})(F_{k-1}(X_k - \hat{X}_{k|k-1}) + G_{k-1}U_{k-1})^T]$

$K_{k|k-1} = F_{k-1} K_{k-1|k-1} F_{k-1}^T + G_{k-1} Q_{k-1} G_{k-1}^T$

Where: $K_{k|k} = var(X_k - \hat{X}_{k|k}) = (I - \Gamma_k H_k) K_{k|k-1} (I - \Gamma_k H_k)^T + \Gamma_k R_k \Gamma_k^T \rightarrow$

$K_{k|k} = (I - \Gamma_k H_k) K_{k|k-1}$

H. Question 8: Apply the results in the model for both observation data generation and Kalman filtering and repeat the simulation scenarios in Part I.

No major differences are seen in the simulations.