Lecture – 40

Kalman Filter Theory

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Outline

- Continuous-time Kalman Filter (CKF)
- Discrete-time Kalman Filter (DKF)
- Continuous-Discrete Kalman Filter (CDKF)
- Extended Kalman Filter (EKF)

Continuous-time Kalman Filter Design for Linear Time Invariant (LTI) Systems

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Problem Statement

System Dynamics: $\dot{X} = AX + BU + GW$ $\quad W(t)$: Process noise vector

Measured Output: $Y = CX + V$ $\quad V(t)$: Sensor noise vector

Assumptions:

(*i*) $X(0) \sim \left(\tilde{X}_0, P_0\right)$, $W(t) \sim \left(0, Q\right)$ and $V(t) \sim \left(0, R\right)$

are "mutually orthogonal" $\left[X(0):\ \text{initial condition for } X\right]$

(*ii*) $W(t)$ and $V(t)$ are uncorrelated white noise

(*iii*) $E\left[W(t)\, W^T(t+\tau)\right] = Q\,\delta(\tau), \quad Q \geq 0$ (psdf)

$E\left[V(t)\, V^T(t+\tau)\right] = R\,\delta(\tau), \quad R > 0$ (pdf)

Problem Statement

Objective:

To obtain an estimate of the state vector $\hat{X}(t)$ using the state dynamics as well as a "sequence of measurements" as accurate as possible.

i.e., to make sure that the error $\tilde{X}(t) \triangleq \left[X(t) - \hat{X}(t) \right]$ becomes very small $\left(\text{ideally } \tilde{X}(t) \rightarrow 0 \right)$ as $t \rightarrow \infty$.

Observer/Estimator/Filter Dynamics

$$\dot{\hat{X}} = A\hat{X} + BU + K_e\left(Y - \hat{Y}\right)$$

where (i) $\hat{X} = E(X)$: Estimate of the state X

(ii) $\hat{Y} = E(Y)$: Estimate of the output Y

$$\begin{aligned} &= E\left(CX + V\right) \\ &= E\left(CX\right) + E\left(V\right) \\ &= CE\left(X\right) \quad \left(\because E(V) = 0\right) \\ &= C\hat{X} \end{aligned}$$

(iii) K_e : Estimator/Filter/Kalman Gain

Problem : How to design K_e ?

Solution: Summary

(i) Initialize $\hat{X}(0)$

(ii) Solve for Riccati matrix P from the Filter ARE:

$$AP + PA^T - PC^T R^{-1} CP + GQG^T = 0$$

(iii) Compute Kalman Gain:

$$K_e = PC^T R^{-1}$$

(iv) Propagate the Filter dynamics:

$$\dot{\hat{X}} = A\hat{X} + BU + K_e\left(Y - C\hat{X}\right)$$

where Y is the measurement vector

Discrete-time Kalman Filter (DKF)

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Model:

$$X_{k+1} = A_k X_k + B_k U_k + G_k W_k$$

$$Y_k = C_k X_k + V_k$$

where W_k and V_k are zero-mean, uncorrelated, Gaussian white noises

$$E\left[W_k W_j^T\right] = Q_k \ \delta_{kj}$$

$$E\left[V_k V_j^T\right] = R_k \ \delta_{kj}$$

$$E\left[V_k W_k^T\right] = 0 \quad \forall k$$

$$\delta_{kj} = \begin{bmatrix} 0 & k \neq j \\ 1 & k = j \end{bmatrix}$$

Estimator: (Predictor-Corrector form)

$$\text{Predictor:} \quad \hat{X}_{k+1}^{-} = A_k \hat{X}_k^{+} + B_k U_k \qquad (1)$$

$$\text{Corrector:} \quad \hat{X}_k^{+} = \hat{X}_k^{-} + K_{e_k} \left[Y_k - C_k \hat{X}_k^{-} \right] \qquad (2)$$

Observer (Recursive) form:

Substituting (2) in (1)

$$\hat{X}_{k+1}^{-} = A_k \hat{X}_k^{-} + B_k U_k + A_k K_{e_k} \left[Y_k - C_k \hat{X}_k^{-} \right]$$

Note: Prediction-Correction form is more popular since its logical, more structured and easy to implement form.

It also leads to a logical extension in Extended Kalman Filter design

Definitions

Error

$$\tilde{X}_k^- \triangleq X_k - \hat{X}_k^- \qquad \tilde{X}_{k+1}^- \triangleq X_{k+1} - \hat{X}_{k+1}^-$$

$$\tilde{X}_k^+ \triangleq X_k - \hat{X}_k^+ \qquad \tilde{X}_{k+1}^+ \triangleq X_{k+1} - \hat{X}_{k+1}^+$$

Error Covariance Matrices

$$P_k^- \triangleq E\left[\tilde{X}_k^- \ \tilde{X}_k^{-T}\right] \qquad P_{k+1}^- \triangleq E\left[\tilde{X}_{k+1}^- \ \tilde{X}_{k+1}^{-\,T}\right]$$

$$P_k^+ \triangleq E\left[\tilde{X}_k^+ \ \tilde{X}_k^{+T}\right] \qquad P_{k+1}^+ \triangleq E\left[\tilde{X}_{k+1}^+ \ \tilde{X}_{k+1}^{+\,T}\right]$$

Objective : To derive expressions for P_{k+1}^-, P_k^+ and K_{e_k}

Expression for P_{k+1}^{-}

$$
\begin{aligned}
\tilde{X}_{k+1}^{-} &= X_{k+1} - \hat{X}_{k+1}^{-} \\
&= \left(A_k X_k + B_k U_k + G_k W_k \right) - \left(A_k \hat{X}_k^{+} + B_k U_k \right) \\
&= A_k \left(X_k - \hat{X}_k^{+} \right) + G_k W_k \\
&= A_k \tilde{X}_k^{+} + G_k W_k
\end{aligned}
$$

Expression for P_{k+1}^{-}

$$P_{k+1}^{-} = E\left[\tilde{X}_{k+1}^{-} \ \tilde{X}_{k+1}^{-\,T}\right]$$

$$= E\left[\left(A_k \tilde{X}_k^{+} + G_k W_k\right)\left(A_k \tilde{X}_k^{+} + G_k W_k\right)^T\right]$$

$$= E\left[A_k \tilde{X}_k^{+} \tilde{X}_k^{+T} A_k^T + G_k W_k \tilde{X}_k^{+T} A_k^T + A_k \tilde{X}_K^{+} W_k^T G_k^T + G_k W_k W_k^T G_k^T\right]$$

$$= A_k \ E\left[\tilde{X}_k^{+} \tilde{X}_k^{+T}\right] A_k^T + G_k \ E\left[W_k \tilde{X}_k^{+T}\right] A_k^T$$

0

$$+ A_k \ E\left[\tilde{X}_k^{+} W_k^T\right] G_k^T + G_k \ E\left[W_k W_k^T\right] G_k^T$$

0

$$\boxed{P_{k+1}^{-} = A_k \ P_k^{+} \ A_k^T + G_k \ Q_k \ G_k^T}$$

$$P_0^{-} = E\left[\tilde{X}_0^{-} \ \tilde{X}_0^{-T}\right]$$

(Note: Only $\tilde{X}_{k+1}$ depends on W_k, not $\tilde{X}_k$; i.e. $\tilde{X}_k^{+}$ and W_k are "orthogonal")

Expression for P_k^+

$$\tilde{X}_k^+ = X_k - \hat{X}_k^+$$

$$= X_k - \left[\hat{X}_k^- + K_{e_k} \left(Y_k - C_k \hat{X}_k^- \right) \right]$$

$$= X_k - \left[\hat{X}_k^- + K_{e_k} \left(C_k X_k + V_k - C_k \hat{X}_k^- \right) \right]$$

$$= \left(I - K_{e_k} C_k \right) X_k - \left(I - K_{e_k} C_k \right) \hat{X}_k^- - K_{e_k} V_k$$

$$= \left(I - K_{e_k} C_k \right) \left(X_k - \hat{X}_k^- \right) - K_{e_k} V_k$$

$$\boxed{\tilde{X}_k^+ = \left(I - K_{e_k} C_k \right) \tilde{X}_k^- - K_{e_k} V_k}$$

Expression for P_k^+

$$P_k^+ = E\left[\tilde{X}_k^+ \ \tilde{X}_k^{+T}\right]$$

$$= E\left[\left[\left(I - K_{e_k} C_k\right)\tilde{X}_k^- - K_{e_k} V_k\right]\left[\left(I - K_{e_k} C_k\right)\tilde{X}_k^- - K_{e_k} V_k\right]^T\right]$$

$$= E\left[\left(I - K_{e_k} C_k\right)\tilde{X}_k^- \tilde{X}_k^{-T}\left(I - K_{e_k} C_k\right)^T\right]$$

$$- E\left[\left(I - K_{e_k} C_k\right)\tilde{X}_k^- V_k^T K_{e_k}^T\right]$$

$$- E\left[K_{e_k} V_k \tilde{X}_k^{-T}\left(I - K_{e_k} C_k\right)^T\right] + E\left[K_{e_k} V_k V_k^T K_{e_k}^T\right]$$

(Note: $\tilde{X}_k^-$ and V_k are "orthogonal")

$$P_k^+ = \left(I - K_{e_k} C_k\right) P_k^- \left(I - K_{e_k} C_k\right)^T + K_{e_k} R_k K_{e_k}^T$$

Expression for K_{e_k} : Problem

Select K_{e_k} such that $Tr\left(P_k^+\right)$ is minimized.

i.e. Minimize $\boxed{J = \dfrac{1}{2} Tr\left(P_k^+\right)}$

with proper selection of K_{e_k}

Expression for K_{e_k} : Solution

$$\frac{\partial J}{\partial K_{e_k}} = 0$$

$$\left(I - K_{e_k} C_k\right) P_k^- C_k^T + K_{e_k} R_k = 0$$

$$K_{e_k}\left[C_k P_k^- C_k^T + R_k\right] = P_k^- C_k^T$$

$$K_{e_k} = P_k^- C_k^T \left[C_k P_k^- C_k^T + R_k\right]^{-1}$$

Simplified Expression for P_k^+

$$
\begin{aligned}
P_k^+ &= \left(I - K_{e_k} C_k\right) P_k^- \left(I - K_{e_k} C_k\right)^T + K_{e_k} R_k K_{e_k}^T \\
&= \left(P_k^- - K_{e_k} C_k P_k^-\right)\left(I - K_{e_k} C_k\right)^T + K_{e_k} R_k K_{e_k}^T \\
&= \left(I - K_{e_k} C_k\right) P_k^- - P_k^- C_k^T K_{e_k}^T + P_k^- C_k^T K_{e_k}^T \\
&= \left(I - K_{e_k} C_k\right) P_k^-
\end{aligned}
$$

Note: Even though this simplification is possible, it is still advisable to use the original expression to avoid numerical problems.

Summary

Model	$X_{k+1} = A_k X_k + B_k U_k + G_k W_k$ $Y_k = C_k X_k + V_k$
Initialization	$\hat{X}(t_0) = \hat{X}_0^-$ $P_0^- = E\left[\tilde{X}_0^- \tilde{X}_0^{-T}\right]$
Gain Computation	$K_{e_k} = P_k^- C_k^T \left[C_k P_k^- C_k^T + R_k\right]^{-1}$

Summary

Updation	$\hat{X}_k^+ = \hat{X}_k^- + K_{e_k}\left[Y_k - C_k \hat{X}_k^-\right]$ $P_k^+ = \left(I - K_{e_k} C_k\right) P_k^- \left(I - K_{e_k} C_k\right)^T + K_{e_k} R_k K_{e_k}^T$ (preferable) $= \left(I - K_{e_k} C_k\right) P_k^-$ (not preferable)
Propagation	$\hat{X}_{k+1}^- = A_k \hat{X}_k^+ + B_k U_k$ $P_{k+1}^- = A_k P_k^+ A_k^T + G_k Q_k G_k^T$

Continuous-Discrete Kalman Filter

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Continuous-Discrete KF

Continuous time model and discrete time measurements

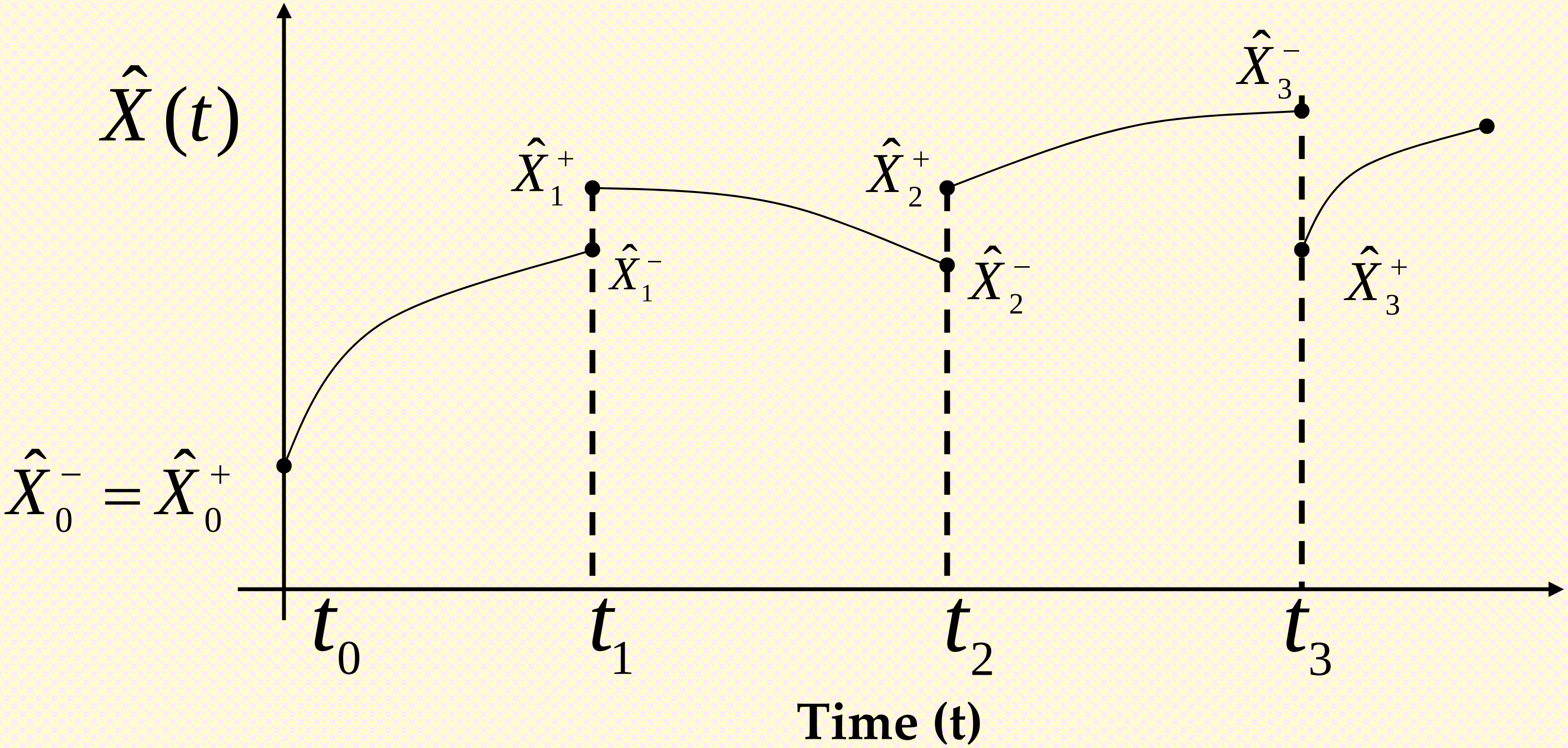
$$\dot{X}(t) = A(t)X(t) + B(t)U(t) + G(t)W(t)$$

$$Y_k = C_k X_k + V_k$$

$$E\left[W(t)W^T(\tau)\right] = Q_k\,\delta(t-\tau)$$

$$E\left[V_k V_j^T\right] = R_k\,\delta_{kj} \qquad \delta_{kj} = \begin{bmatrix} 0 & k \neq j \\ 1 & k = j \end{bmatrix}$$

Mechanism



$\hat{X}(t)$
$\hat{X}_0^- = \hat{X}_0^+$
$\hat{X}_1^+$
$\hat{X}_1^-$
$\hat{X}_2^+$
$\hat{X}_2^-$
$\hat{X}_3^-$
$\hat{X}_3^+$
t_0
t_1
t_2
t_3
Time (t)

Principle

Propagate the state-estimate model forward from t_k to t_{k+1} using the initial condition $\hat{X}_k^+$ *i.e.*, $\hat{X}_k^+ \rightarrow \hat{X}_{k+1}^-$

Correct the value $\hat{X}_{k+1}^-$ to $\hat{X}_{k+1}^+$ using the measurement vector Y_{k+1}

Measurement is available only at discrete time-steps. Hence, the continuum time propagation model DOES NOT involve any measurement information. This leads to:

$$\dot{P}(t) = AP + PA^T + GQG^T$$

Expression for $\dot{P}$

$$\dot{X} = AX + BU + GW$$
$$\dot{\hat{X}} = A\hat{X} + BU$$

$$\Rightarrow$$

$$\dot{\tilde{X}} = \dot{X} - \dot{\hat{X}} = A\tilde{X} + GW$$

$$\tilde{X}(t) = \varphi(t,t_0)\tilde{X}_0 + \int_0^t \varphi(t,t_0)G(\tau)W(\tau)\,d\tau$$

$$R_{W\tilde{X}} = E\left[\int_0^t W(t)\,W^T(\tau)G(\tau)\varphi(t,\tau)\,d\tau\right]$$

$$= \int_0^t Q\,\delta(t-\tau)G^T(\tau)\varphi(t,\tau)\,d\tau = \tfrac{1}{2}Q\,G^T$$

Expression for $\dot{P}$

$$\dot{P} = E\left[\dot{\tilde{X}}\,\tilde{X}^T + \tilde{X}\,\dot{\tilde{X}}^T\right] = E\left[\dot{\tilde{X}}\,\tilde{X}^T\right] + \left(E\left[\dot{\tilde{X}}\,\tilde{X}^T\right]\right)^T$$

$$E\left[\dot{\tilde{X}}\,\tilde{X}^T\right] = E\left[\left(A\tilde{X} + GW\right)\tilde{X}^T\right]$$

$$= AE\left[\tilde{X}\,\tilde{X}^T\right] + G\,E\left[W\,\tilde{X}^T\right]$$

$$= AP + \tfrac{1}{2}GQG^T$$

$$\dot{P} = \left(AP + \tfrac{1}{2}GQG^T\right) + \left(AP + \tfrac{1}{2}GQG^T\right)^T$$

$$\boxed{\dot{P} = AP + PA^T + GQG^T}$$

Summary

Model	$\dot{X}(t)=A(t)X(t)+B(t)U(t)+G(t)W(t)$ $Y_k=C_kX_k+V_k$
Initialization	$\hat{X}(t_0)=\hat{X}_0$ $P_0^-=E\left[\tilde{X}(t_0)\tilde{X}^T(t_0)\right]$
Gain Computation	$K_{e_k}=P_k^-C_k^T\left[C_kP_k^-C_k^T+R_k\right]^{-1}$

Summary

Updation	$\hat{X}_k^+ = \hat{X}_k^- + K_{e_k}\left[Y_k - C_k \hat{X}_k^-\right]$ $P_k^+ = \left(I - K_{e_k} C_k\right) P_k^- \left(I - K_{e_k} C_k\right)^T + K_{e_k} R_k K_{e_k}^T$ (preferable) $= \left(I - K_{e_k} C_k\right) P_k^-$ (not preferable)
Propagation (using high accuracy numerical integration)	$\dot{\hat{X}} = A\hat{X} + BU$ $\dot{P}(t) = AP + PA^T + GQG^T$

Note:

$\dot{P}(t)$ expression is a continuous time Lyapunov Equation (its linear in $P(t)$)

Continuous-discrete Kalman filter facilitates the usage of non-uniform Δt

Use $\dot{\hat{X}}(t)$ and $\dot{P}(t)$ expressions to propagate $\hat{X}_k^+ \rightarrow \hat{X}_{k+1}^-$ and $\hat{P}_k^+ \rightarrow \hat{P}_{k+1}^-$ respectively.

Extended Kalman Filter (EKF)

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Facts to Remember

Nonlinear estimation problems are considerably more difficult than the linear problem in general. (EKF is just an idea…not a cure for everything !)

The problem with nonlinear systems is that a Gaussian input does not necessarily produce a Gaussian output (unlike linear case)

The EKF even though not truly `optimum', has been successfully applied in many nonlinear systems over the decades

The fundamental assumption in EKF design is that the true state $X(t)$ is sufficiently close to the estimated state $\hat{X}(t)$ at all time, and hence the error dynamics can be represented fairly accurately by the linearized system about $\hat{X}(t)$

Continuous-Discrete EKF

Dr. Radhakant Padhi
Asst. Professor
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Continuous-Discrete EKF

Motivation: System dynamics is a continuous-time, whereas measurements are available only at discrete interval of time.

Strategy:

Without the availability of measurement, propagate the state and co-variance dynamics from $\hat{X}_k^+ \rightarrow \hat{X}_{k+1}^-$ and $P_k^+ \rightarrow P_{k+1}^-$ respectively, using the "nonlinear system dynamics" and linear co-variance dynamics.

As soon as the measurement is available, update $\hat{X}_k^- \rightarrow \hat{X}_k^+$ and $P_k^- \rightarrow P_k^+$ respectively.

Summary: Continuous-Discrete EKF

Model	$\dot{X}(t) = f(X, U, t) + G(t)W(t)$ $Y_k = h(X_k) + V_k$
Initialization	$\hat{X}^-(t_0) = X_0$ $P_0^- = E\left[\tilde{X}^-(t_0)\,\tilde{X}^{-T}(t_0)\right]$
Gain Computation	$K_{e_k}(t) = P_k^- C_k^{-T}\left[C_k^- P_k^- C_k^{-T} + R\right]^{-1}$ where, $C_k^- = \left[\dfrac{\partial h}{\partial X}\right]_{\hat{X}_k^-}$

Summary: Continuous-Discrete EKF

Updation	$\hat{X}_k^+ = \hat{X}_k^- + K_{e_k}\left[Y_k - h\left(\hat{X}_k^-\right)\right]$ $P_k^+ = \left(I - K_{e_k} C_k\right) P_k^- \left(I - K_{e_k} C_k\right)^T + K_{e_k} R_k K_{e_k}^T$ (preferable) $= \left(I - K_{e_k} C_k\right) P_k^-$ (not preferable)
Propagation	$\dot{\hat{X}}(t) = f(\hat{X}, U, t); \quad \hat{X}_k^+ \rightarrow \hat{X}_{k+1}^-$ $\dot{P}(t) = AP + PA^T + GQG^T; \ \hat{P}_k^+ \rightarrow \hat{P}_{k+1}^-$ where $A(t) = \left[\dfrac{\partial f}{\partial X}\right]_{\hat{X}(t)}$

Iterated EKF

One way of improving the performance of EKF is to apply local iterations to repeatedly calculate $\hat{X}_k^+, \hat{P}_k^+$ and K_{e_k}, each time by linearising about the most recent estimate. This approach is known as `Iterative EKF'.

Note: One can proceed with a fixed number of iterations.

Recommendations/Issues in EKF

- **Design parameter selection:**
  - Fix R based on the sensor characteristics
  - Select P_0 to be “sufficiently high”
  - Tune Q until obtaining satisfactory results
- **The filter should run sufficiently ahead of time prior to its usage, so that the error stabilizes before its actual usage (else, initial error can be very large and the associated control can destabilize the closed loop system)**
- **Keep the measurement equation linear wherever possible**
- **Care should be taken to avoid numerical ill-conditioning. Methods are available to address this issue (see Crassidis and Jenkins book).**

Recommendations/Issues in EKF

- **Care should be taken to eliminate the `outliers'. For e.g., if the measurement output is too far away from the predicted output, then it can be treated as an outlier.**
- **EKF is `fragile', i.e., only a narrow band of design variables P_0, R, and Q exists for its success. Hence, the tuning is necessary for any given application. Hence tuning process should be done very carefully.**
- **Checks for Consistency of Kalman Filter :**
  - **Sigma-bound test**
  - **Normalized Error Square (NES) test**
  - **Normalized Mean Error (NME) test**
  - **Autocorrelation Test**
  - **Cramer-Rao Inequality (gives a "lower bound" on error)**

Limitations of EKF

- Linearization can introduce significant error
- No general convergence guarantee
- Works in general; but in some cases its performance can be surprisingly bad
- Unreliable for colour noise (shaping filter philosophy need not hold good in general)

Beyond EKF

- **Need**
  - Nonlinear systems
  - Non Gaussian noise
  - Correlated noise
  - Colour noise
- **Characteristics of Such Filters**
  - Are often approximate
  - Sacrifices theoretical accuracy in favour of practical constraints and considerations like robustness, adaptation, numerical feasibility
  - Attempt to cover the limitations of EKF

References: Books

- J. L. Crassidis and J. L. Junkins, **Optimal Estimation of Dynamic Systems**, *CRC Press, 2004.*
- D. Simon, **Optimal State Estimation**, Wiley, 2006.
- P. Zarchan and H. Musoff, **Fundamentals of Kalman Filtering: A Practical Approach**, *AIAA, 2005.*
- F. L. Lewis, L. Xie and D. Popa, **Optimal and Robust Estimation**, *CRC Press, 2007.*
- A. Gelb, **Applied Optimal Estimation**, *MIT Press, 1974.*

Thanks for the Attention….!