Lecture – 30

Kalman Filter Design – III

Prof. Radhakant Padhi
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Continuous-Discrete Kalman Filter

Prof. Radhakant Padhi
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore
**Continuous-Discrete KF**

Continuous time model and discrete time measurements

\[
\dot{X}(t) = A(t)X(t) + B(t)U(t) + G(t)W(t) \\
Y_k = C_k X_k + V_k
\]

\[
E\left[ W(t)W^T(\tau) \right] = Q_k \delta(t - \tau) \\
E\left[ V_k V_j^T \right] = R_k \delta_{kj}
\]

\[
\delta_{kj} = \begin{cases} 
0 & \text{if } k \neq j \\
1 & \text{if } k = j 
\end{cases}
\]

**Mechanism**

\[
\hat{X}(t) = \hat{X}_0 \\
\hat{X}_0^+ = \hat{X}_0^-
\]

\[
t_0 \quad t_1 \quad t_2 \quad t_3
\]

\[
\hat{X}_1^- \quad \hat{X}_2^- \quad \hat{X}_3^-
\]

\[
\hat{X}_1^+ \quad \hat{X}_2^+ \quad \hat{X}_3^+
\]
**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 \( \dot{P} \)**

\[
\begin{align*}
\dot{X} &= AX + BU + GW \\
\hat{X} &= AX + BU
\end{align*}
\Rightarrow
\begin{align*}
\dot{\hat{X}} &= \hat{X} - \dot{X} \\
&= A\hat{X} + GW
\end{align*}
\]

\[
\begin{align*}
\tilde{X}(t) &= \phi(t, t_0)\tilde{X}_0 + \int_0^t \phi(t, \tau)G(\tau)W(\tau) d\tau \\
R_{WX} &= E \left[ \int_0^t W(t) W^T(\tau) G(\tau) \phi(t, \tau) \ d\tau \right] \\
&= \int_0^t Q \delta(t-\tau) G^T(\tau) \phi(t, \tau) \ d\tau = \frac{1}{2} Q G^T
\end{align*}
\]
Expression for $\dot{P}$

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

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

\[
= AE \left[ \dot{X} \dot{X}^T \right] + GE \left[ W \dot{X}^T \right]
\]

\[
= AP + \frac{1}{2} GQG^T
\]

\[
\dot{P} = (AP + \frac{1}{2} GQG^T) + (AP + \frac{1}{2} GQG^T)^T
\]

\[
\dot{P} = AP + PA^T + GQG^T
\]

---

Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$\dot{X}(t) = A(t)X(t) + B(t)U(t) + G(t)W(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Y_k = C_k X_k + V_k$</td>
</tr>
</tbody>
</table>

| Initialization | $\dot{X}(t_0) = \dot{X}_0$ |
|               | $P_0^- = E \left[ \dot{\dot{X}}(t_0) \dot{X}^T(t_0) \right]$ |

| Gain Computation | $K_{\epsilon_k} = P_k^- C_k^T \left[ C_k P_k^- C_k^T + R_k \right]^{-1}$ |
## Summary

<table>
<thead>
<tr>
<th>Updation</th>
<th>$\hat{X}<em>i = \hat{X}</em>{i-1} + K_{i-1} \left[ Y_i - C_i \hat{X}_{i-1} \right]$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P_i^+ = (I - K_{i-1} C_i) P_i^- \left( I - K_{i-1} C_i \right)^T + K_{i-1} R_i K_{i-1}^T$ (preferable)</td>
</tr>
<tr>
<td></td>
<td>$= (I - K_{i-1} C_i) P_i^-$ (not preferable)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Propagation</th>
<th>Continuous-discrete Kalman filter facilitates the usage of non-uniform $\Delta t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Use $\dot{X}(t)$ and $\dot{P}(t)$ expressions to propagate $\hat{X}<em>k^+ \rightarrow \hat{X}</em>{k+1}^-$ and $\hat{P}<em>k^+ \rightarrow \hat{P}</em>{k+1}^-$ respectively.</td>
</tr>
</tbody>
</table>

## 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 $\Delta t$
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-Continuous EKF

Prof. Radhakant Padhi
Dept. of Aerospace Engineering
Indian Institute of Science - Bangalore

Continuous-Continuous Formulation

(Assumption: Measurement is continuously available)

\[ \dot{X}(t) = f(X, U, t) + G(t)W(t) \]
\[ Y = h(X, t) + V(t) \]

where \( f \) and \( h \) are "continuously differentiable" and \( W(t), V(t) \) are uncorrelated, zero mean Gaussian white noises.
**Formulation**

\[ X(t) \triangleq \hat{X}(t) + \tilde{X}(t) \]

Taylor series expansion and neglecting HOTs

\[ f(X,U,t) = f(\hat{X},U,t) + \left[ \frac{\partial f}{\partial X} \right] \tilde{X} \]

\[ = f(\hat{X},U,t) + \left[ \frac{\partial f}{\partial X} \right] \left( X - \hat{X} \right) \]

Similarly,

\[ h(X,t) = h(\hat{X},t) + \left[ \frac{\partial h}{\partial X} \right] \left( X - \hat{X} \right) \]

**Formulation**

\[ E[f(X,U,t)] = E[f(\hat{X},U,t)] + \left[ \frac{\partial f}{\partial X} \right] \left( E[X] - \hat{X} \right) \]

\[ = f(\hat{X},U,t) + \left[ \frac{\partial f}{\partial X} \right] \left( \hat{X} - \hat{X} \right) \]

\[ = f(\hat{X},U,t) \]

Similarly, \[ E[h(X,t)] = h(\hat{X},t) \]

Observer dynamics:

\[ \dot{\hat{X}}(t) = f(\hat{X},U,t) + K_c(t) \left[ Y - h(\hat{X},t) \right] \]

\[ \hat{Y} = h(\hat{X},t) \]
Error Dynamics

\[ \dot{X}(t) \triangleq X(t) - \ddot{X}(t) \]

\[ \dot{X}(t) = \{f(X,U,t) + GW\} - \{f(\ddot{X},U,t) + K_e\left[h(X,t) + V - h(\dddot{X},t)\right]\} \]

\[ = \left[f(X,U,t) - f(\ddot{X},U,t)\right] - K_e\left[h(X,t) - h(\dddot{X},t)\right] + GW - K_e V \]

\[ = \left[\frac{\partial f}{\partial X}\right]_\ddot{X} - K_e\left[\frac{\partial h}{\partial X}\right]_\dddot{X} \hat{x} + GW - K_e V \]

This error dynamics is EXACTLY SAME as the error dynamics derived for the time-varying linear system case. Hence, the dynamics for the error Covariance matrix is given by

\[ \dot{P}(t) = AP + PA^T - PC^T R^{-1} CP + GQG^T \]

where, \( A(t) = \left[\frac{\partial f}{\partial X}\right]_\ddot{X} \); \( C(t) = \left[\frac{\partial h}{\partial X}\right]_\dddot{X} \)
## Summary (Continuous EKF)

| Model       | \( \dot{X}(t) = f(X,U,t) + G(t)W(t) \)  
|            | \( Y = h(X,t) + V(t) \) |
| Initialization | \( \dot{X}(t_0) = \hat{X}_0 \)  
|            | \( P_0 = E[\tilde{X}(t_0)\tilde{X}^T(t_0)] \) |
| Gain Computation | \( K_e(t) = P(t)C^T(t)R^{-1}(t) \) |

## Summary

| Propagation | \( \dot{X}(t) = f(\hat{X},U,t) + K_e(t)[Y - h(\hat{X},t)] \)  
|            | \( \dot{P}(t) = AP + PA^T - PC^T \)  
|            | \( + PC^T R^{-1}CP + GQG^T \) |
| where, \( A(t) = \left[ \frac{\partial f}{\partial X} \right]_{\hat{X}} \); \( C(t) = \left[ \frac{\partial h}{\partial X} \right]_{\hat{X}} \) |
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** | \( \dot{X}(t) = f(X, U, t) + G(t)W(t) \)  
|           | \( Y = h(X_k) + V_k \) |
| **Initialization** | \( \dot{X}^{-}(t_0) = X_0 \)  
|           | \( P_0^- = E \left[ \dot{X}^{-}(t_0) \dot{X}^{-T}(t_0) \right] \) |
| **Gain Computation** | \( K_{\epsilon_k}(t) = P_k^- C_k^{-T} \left[ C_k^- P_k^- C_k^{-T} + R \right]^{-1} \)  
|           | where, \( C_k^- = \left[ \frac{\partial h}{\partial X} \right] \dot{X}_k^- \) |

### Summary: Continuous-Discrete EKF

| **Updation** | \( \dot{X}_i^+ = \dot{X}_i^- + K_{\epsilon_i} \left[ Y_i - h(\dot{X}_i^-) \right] \)  
|              | \( P_i^+ = \left( I - K_{\epsilon_i} C_i \right) P_i^- \left( I - K_{\epsilon_i} C_i \right)^T + K_{\epsilon_i} R_{\epsilon_i} K_{\epsilon_i}^T \) \( \text{preferable} \)  
|              | \( = \left( I - K_{\epsilon_i} C_i \right) P_i^- \) \( \text{(not preferable)} \) |

| **Propagation** | \( \dot{X}(t) = f(\dot{X}, U, t) \);  
|                | \( \dot{X}_k^+ \rightarrow \dot{X}_{k+1} \)  
|                | \( \dot{P}(t) = AP + PA^T + GQG^T \); \( \dot{P}_k^+ \rightarrow \dot{P}_{k+1} \)  
|                | where \( A(t) = \left[ \frac{\partial f}{\partial X} \right] \dot{X}(t) \) |
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 $K_{k_1}$, 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.

Linearized Kalman Filter (LKF)

This approach involves linearization about a nominal state trajectory $\bar{X}(t)$ (which is selected a priori), instead of the current estimate $\hat{X}(t)$. In such a situation,

$$f(X,U,t) = f(\bar{X},U,t) + A(t)[X - \bar{X}]$$
$$h(X,U,t) = h(\bar{X},U,t) + C(t)[X - \bar{X}]$$

where, $A(t) = \left[ \frac{\partial f}{\partial X} \right]_{\bar{X}}$; $C(t) = \left[ \frac{\partial h}{\partial X} \right]_{\bar{X}}$
**Linearized Kalman Filter (LKF)**

\[
E[f(X,U,t)] = f(\bar{X},U,t) + A(t)\left[\hat{X} - \bar{X}\right]
\]

\[
E[h(X,U,t)] = h(\bar{X},U,t) + C(t)\left[\hat{X} - \bar{X}\right]
\]

Hence, the LKF has the following structure

\[
\dot{\hat{X}}(t) = f(\bar{X},U,t) + A(t)\left[\hat{X} - \bar{X}\right] + K_e(t)\left[Y - h(\bar{X},U,t) - C(t)\left[\hat{X} - \bar{X}\right]\right]
\]

\[
\hat{Y} = h(\bar{X},U,t) + C(t)\left[\hat{X} - \bar{X}\right]
\]

**Linearized Kalman Filter (LKF)**

The Co-variance matrix can be derived to be

\[
\dot{P}(t) = AP + PA^T - PC^T R^{-1} CP + GQG^T
\]

where, \(A(t) = \left[\frac{\partial f}{\partial X}\right]_{\bar{X}(t)}\) ; \(C(t) = \left[\frac{\partial h}{\partial X}\right]_{\bar{X}(t)}\)

\(\bar{X}(t)\) is the nominal state
Comments

In general, LKF is less accurate than EKF, since $X(t)$ is usually not close to $\hat{X}(t)$ as $\hat{X}(t)$ is.

LKF can, however, lead to the a priori computation of $K_e(t)$, which can then be stored and used with less online computation.

To avoid the large initial chattering of EKF, one can start with LKF and then switch to EKF. However, once it is often possible to start the EKF ahead of time, this necessity normally does not arise.

The philosophy of local iterations can also be implemented with LKF, leading to `Iterated LKF'.

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, it can be treated as an outlier. Data-rejection methods are available.
- EKF is 'fragile', i.e., only a narrow band of design variables $P$, $R$, and $Q$ exists for its success. Hence, tuning is necessary for any given application and the 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)

Difficulties in Practice

- Computer round-off errors
- Unchecked error propagation
- Asymmetry of covariance matrix: A symptom of numerical degradation
- Solution of Riccati equation
  - Square-root filtering: Solution of Riccati equation via Cholesky factorization
- Large initial errors: Information filtering
Point 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 $\hat{X}(t)$ 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 $\hat{X}(t)$.

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 (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

---

**Unscented Kalman Filter (UKF)**

**Prof. Radhakant Padhi**

Dept. of Aerospace Engineering

Indian Institute of Science - Bangalore
**Motivation**

- Tuning difficulties in EKF
- Change of characteristics of noise PDF through nonlinear transformation
- Gaussian distribution properties can be propagated easily rather than covariance matrix
- Approximation of higher order moment of the distribution after transformation
A Simple Example

A comparison of the exact, linearized, and unscented mean and covariance of 300 randomly generated points with $r$ uniformly distributed between $0.01$ and $0.35$ radian $\theta$ uniformly distributed between $\pm 0.01$ and $\tilde{\theta}$ uniformly distributed between $\pm 0.35$ radians.


System Dynamics

\[
x_{k+1} = f(x_k, u_k, t_k) + w_k
\]
\[
y_k = h(x_k, t_k) + v_k
\]
\[
w_k \sim (0, Q_k)
\]
\[
v_k \sim (0, R_k)
\]

Sigma Point Approach

1. A set of weighted samples (sigma-points) are deterministically calculated using the mean and square-root decomposition of the covariance matrix of the prior random variable. As a minimal requirement the sigma-point set must completely capture the first and second order moments of the prior random variable. Higher order moments can be captured, if so desired, at the cost of using more sigma-points.

2. The sigma-points are propagated through the true nonlinear function using functional evaluations alone, i.e., no analytical derivatives are used, in order to generate a posterior sigma-point set.

3. The posterior statistics are calculated (approximated) using tractable functions of the the propagated sigma-points and weights. Typically these take on the form of simple weighted sample mean and covariance calculations of the posterior sigma-points.


UKEF Implementation: Propagation of State Estimation and Covariance

Initialization:

\[
\hat{x}_0^+ = E(x_0) \\
P_0^+ = E [(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T]
\]

Selection of \(\sigma\)-points:

\[
\hat{x}_{k-1}^{(i)} = \hat{x}_{k-1}^+ + \tilde{x}^{(i)} \\
\tilde{x}^{(i)} = \left(\sqrt{nP_{k-1}^+}\right)_i^T \\
\tilde{x}^{(n+i)} = -\left(\sqrt{nP_{k-1}^+}\right)_i^T \\
\]

Note:

\[
\left(\sqrt{A}\right)^T \sqrt{A} = A
\]
UKF Implementation: Propagation of State Estimation and Covariance

Transformation of $\sigma$-points:

$$\hat{x}_k^{(i)} = f(\hat{x}_{k-1}^{(i)}, u_k, t_k)$$

Calculate new mean and covariance:

$$\hat{x}_k^- = \frac{1}{2n} \sum_{i=1}^{2n} \hat{x}_k^{(i)}$$

$$P_k^- = \frac{1}{2n} \sum_{i=1}^{2n} (\hat{x}_k^{(i)} - \hat{x}_k^-) (\hat{x}_k^{(i)} - \hat{x}_k^-)^T + Q_{k-1}$$

UKF Implementation: Measurement Update

Selection of $\sigma$-points:

$$\tilde{x}_k^{(i)} = \hat{x}_k^- + \tilde{x}^{(i)} \quad i = 1, \ldots, 2n$$

$$\tilde{x}^{(i)} = \left( \sqrt{nP_k^-} \right)^T \quad i = 1, \ldots, n$$

$$\tilde{x}^{(n+i)} = -\left( \sqrt{nP_k^-} \right)^T \quad i = 1, \ldots, n$$

Note: previously generated sigma points can also be used. However, that'll lead to some performance degradation.
UKF Implementation: Measurement Update

Transformation of σ-points:

\[ \hat{y}_k^{(i)} = h(\hat{x}_k^{(i)}, t_k) \]

Predicted measurements

\[ \hat{y}_k = \frac{1}{2n} \sum_{i=1}^{2n} \hat{y}_k^{(i)} \]
Advantage vs. Computational Effort


References: Books

References: Publications


Thanks for the Attention....!!