LTV Hybrid Kalman Filter

September 16, 2019

1 Discrete Kalman Filter

A generic discrete linear time-varying (LTV) system can be written in the form

$$x_k = A_k x_{k-1} + B_k u_k + G\mathbf{w}_k \tag{1}$$
$$y_k = C_k x + \mathbf{v}_k \tag{2}$$

with $x \in \mathbb{R}^n$ as the state vector, $u \in \mathbb{R}^m$ as the input vector, $y \in \mathbb{R}^s$ as the output vector, $A \in \mathbb{R}^{nxn}$ as the system matrix, $B \in \mathbb{R}^{nxm}$ as the input matrix and $C \in \mathbb{R}^{sxn}$ as the observation matrix. $\mathbf{v} \sim \mathcal{N}(0, R)$ and $\mathbf{w} \sim \mathcal{N}(0, Q)$ are multivariate white noise processes that represent measurement and process noise, and G is the constant matrix that defines how the process noise $\mathbf{w}$ acts on the system.

The Kalman Filter is a recursive estimator used to estimate the state x based off of previous states and measurements. It uses Bayesian inference to maximize the probability of the estimate of x given the measurement y, and it is represented by two variables

- $\hat{x}_{k|j}$ The expected value, estimate, of x at time k given measurements up to and including time j.
- $P_{k|j}$ The estimate covariance at time k given measurements up to and including time j.

When $k > j$, the probability of $x_{k|j}$,$P(x)$, is referred to the prior probability. When a measurement is received, the probability of y given x, $P(y|x)$, is constructed and referred to the likelihood. The objective of the Kalman Filter is to maximize the posterior probability or the probability of x given y, $P(y|x)$, in two phases: a prediction phase and an update phase.

The prediction phase propagates the state dynamics and the error covariance.

$$\hat{x}_{k|j} = A_k \hat{x}_{k-1|j} + B_k u_k \tag{3}$$
$$P_{k|j} = A_k P_{k-1|j} A_k^\top + G Q_k G^\top \tag{4}$$

note that equation (4)is a Discrete Riccati Equation with P being positive definite and Q semi-positive definite.

The update phase adjusts the state estimate and the estimate covariance to maximize the posterior probability. At the end of this update step $j = k$.

$$K_k = P_k C_k^\top \left(C_k P_k C_k^\top + R_k\right)^{-1} \tag{5}$$
$$\hat{x}_{k|k=j} = \hat{x}_{k|j} + K_k \left(y_k - C_k \hat{x}_{k|j}\right) \tag{6}$$
$$P_{k|j} = \left(I - K_k C_k\right) P_{k|j} \tag{7}$$

2 Continuous Kalman Filter

A generic continuous linear time-varying (LTV) system can be written in the form

$$\dot{x} = A(t)\,x + B(t)\,u + G\mathbf{w}(t) \tag{8}$$
$$y = C(t)\,x + \mathbf{v}(t) \tag{9}$$

where 8and 9are the continuous version of (1)and (2). The update phase for the continuous Kalman filter is the same as the discrete Kalman filter, but the prediction phases are different.

The prediction phase for the continuous Kalman filter is

$$\begin{aligned}\dot{\hat{x}}(t) &= A(t)\hat{x}(t) + B(t)u(t) \quad (10)\\ \Rightarrow \hat{x}_{k|j} &= \hat{x}(t)\\ \dot{P}(t) &= A(t)P(t) + P(t)A^{\top}(t) + Q(t) \quad (11)\\ \Rightarrow P_{k|j} &= P(t)\end{aligned}$$

3 Continuous to Discrete Kalman Filter

Sometimes when numerically integrating P using (11), P can diverge from being positive definite. Also, $P \in \mathbb{R}^{nxn}$ and can become very large when the number of states is very large. Numerically integrating P with many elements can be slow and inaccurate. The solution is to transform (11)to (4).

We begin this process by noting that

$$\dot{x} = A(t)x + B(t)u + G\mathbf{w}(t) \quad (12)$$

has the solution

$$x(t+\delta) = \Phi(t+\delta,t)x(t) + \int_t^{t+\delta}\Phi(t+\delta,\tau)B(\tau)u(\tau)d\tau + \int_t^{t+\delta}\Phi(t+\delta,\tau)G\mathbf{w}(\tau)d\tau \quad (13)$$

where $\Phi(t+\delta,t)$ is the state transition matrix and defined as

$$\begin{aligned}\Phi(t,t_0) = I + \int_{t_0}^{t}A(s_1)ds_1 + \int_{t_0}^{t}A(s_1)\int_{t_0}^{s_1}A(s_2)ds_2ds_1 + \\ = \int_{t_0}^{t}A(s_1)\int_{t_0}^{s_1}A(s_2)\int_{t_0}^{s_2}A(s_2)ds_3ds_2ds_1\ldots\end{aligned}$$

. Assume that δ is the time step of the system and that the input $u(t)$ is constant during that time step. Then (13) can be simplified as

$$x(t+\delta) = \Phi(t+\delta,t)x(t) + \int_t^{t+\delta}\Phi(t+\delta,\tau)B(\tau)ud\tau + \int_t^{t+\delta}\Phi(t+\delta,\tau)G\mathbf{w}(\tau)d\tau$$

We can see that the discrete form of the system matrix $A(t)$ is

$$A_k = \Phi(t+\delta,t) \quad (14)$$

, that the discrete form of the input matrix $B(t)$ is

$$B_k = \int_t^{t+\delta}\Phi(t+\delta,\tau)B(\tau) \quad (15)$$

, and that the discrete form of the noise $\mathbf{w}(t)$ is

$$\mathbf{w}_k = \int_t^{t+\delta}\Phi(t+\delta,\tau)G\mathbf{w}(\tau)d\tau \quad (16)$$

The covariance of $\mathbf{w}_k$can be calculated from the continuous white noise process $\mathbf{w}(t)$. First we note a few properties

$$\begin{aligned}E\left[\mathbf{w}(\tau)\mathbf{w}^{\top}(\alpha)\right] &= Q(\tau)\delta(\tau-\alpha) \quad (17)\\ E\left[\mathbf{w}_k\mathbf{w}_k^{\top}\right] &= Q_k \quad (18)\end{aligned}$$

Q_k can now be solved in terms of $Q(t)$ by substituting (16)into (18) and using (17).

$$\begin{aligned}
Q_k &= E\left[\mathbf{w}_k\mathbf{w}_k^\top\right] \\
&= E\left[\int_{k-1}^{k}\Phi(k,\tau)\,G\mathbf{w}(\tau)\left(\int_{k-1}^{k}\Phi(k,\alpha)\,G\mathbf{w}(\alpha)\,d\alpha\right)^\top d\tau\right] \\
&= E\left[\int_{k-1}^{k}\int_{k-1}^{k}\Phi(k,\tau)\,G\mathbf{w}(\tau)\,\mathbf{w}^\top G^\top(\alpha)\,\Phi(k,\alpha)^\top d\tau d\alpha\right] \\
&= \int_{k-1}^{k}\int_{k-1}^{k}\Phi(k,\tau)\,GE\left[\mathbf{w}(\tau)\,\mathbf{w}^\top(\alpha)\right]G^\top\Phi(k,\alpha)^\top d\tau d\alpha \\
&= \int_{k-1}^{k}\int_{k-1}^{k}\Phi(k,\tau)\,GQ(\tau)\,\delta(\tau-\alpha)\,G^\top\Phi(k,\alpha)^\top d\tau d\alpha \\
&= \int_{k-1}^{k}\Phi(k,\tau)\,GQ(\tau)\,G^\top\Phi(k,\tau)^\top d\tau \qquad (19)
\end{aligned}$$

In the case of linear time-invariant systems, we can make sever simplification.

3.1 LTI

The general from for Linear Time-Invariant (LTI) systems is

$$\begin{aligned}
\dot{x} &= Ax(t)+Bu(t)+G\mathbf{w}(t) \qquad (20)\\
y &= Cx(t)+\mathbf{v}(t) \qquad (21)
\end{aligned}$$

where the matrices A, B and C are constant, and the noise covariance matrices Q and R are also constants.

For Linear Time-Invariant (LTI) systems, the state transition matrix is

$$\begin{aligned}
\Phi(t,t_0) &= \exp^{A(t-t_0)} \qquad (22)\\
&= \sum_{k=0}^{\infty}\frac{(t-t_0)}{k!}A^k \\
&= \sum_{k=0}^{\infty}\frac{\delta}{k!}A^k
\end{aligned}$$

which can be approximated as

$$\Phi(t,t_0)\approx I+A\delta+\frac{(A\delta)^2}{2}+\frac{(A\delta)^3}{6}$$

using (14)

$$\begin{aligned}
A_k = \Phi(t+\delta,t) &= \exp^{A(\delta)} \\
&\approx I+A\delta+\frac{(A\delta)^2}{2}+\frac{(A\delta)^3}{6}
\end{aligned}$$

The discrete input matrix can be simplified to

$$\begin{aligned}
B_k &= \int_{t}^{t+\delta}\exp^{A(t+\delta-\tau)}B d\tau \\
&= A^{-1}\left(\exp^{A\delta}-I\right)B
\end{aligned}$$

which is not always helpful when A is not invertible. In the case that A is not invertible, we can approximate it. But first we use (22) to note that

$$\begin{aligned}
B_k &= A^{-1}\left(\exp^{A\delta} - I\right)B \\
&= A^{-1}\left(I + A\delta + \frac{A^2\delta^2}{2} + \cdots + \frac{A^{\infty}(\tau)^{\infty}}{\infty!} - I\right)B \\
&= A^{-1}\left(A\delta + \frac{A^2\delta^2}{2} + \frac{A^3\delta^3}{6} + \cdots + \frac{A^{\infty}(\tau)^{\infty}}{\infty!}\right)B \\
&= \left(\delta + \frac{A\delta^2}{2} + \frac{A^2\delta^3}{6} + \cdots + \frac{A^{\infty-1}(\tau)^{\infty}}{\infty!}\right)B \\
&= \left(2I + A\delta + \frac{A^2\delta^2}{3} + \cdots + \frac{A^{\infty}(\tau)^{\infty}}{\infty!/2}\right)\frac{\delta}{2}B
\end{aligned}$$

by ignoring higher order terms, we get

$$B_k = \left(2I + A\delta + \frac{A^2\delta^2}{3} + \frac{A^3\delta^3}{12}\right)\frac{\delta}{2}B$$

If we assume that the process noise is constant during the interval δ, then (16) can be simplified as

$$\begin{aligned}
\mathbf{w}_k &= \left(\int_t^{t+\delta} \Phi(t+\delta,\tau)\, G d\tau\right)\mathbf{w} \\
&= G_k\mathbf{w}
\end{aligned}$$

where

$$G_k = \int_t^{t+\delta} \Phi(t+\delta,\tau)\, G d\tau$$

. In similar manner as the input matrix, the system noise matrix G_k becomes

$$\begin{aligned}
G_k &= A^{-1}\left(\exp^{A\delta} - I\right)G \\
&\approx \left(2I + A\delta + \frac{A^2\delta^2}{3} + \frac{A^3\delta^3}{12}\right)\frac{\delta}{2}G
\end{aligned}$$

The process noise covariance can be simplified by substituting (22) into (19)

$$Q_k = \int_t^{t+\delta} \exp^{A(t+\delta-\tau)} GQG^{\top} \exp^{A^{\top}(t+\delta-\tau)} d\tau \tag{23}$$

which can be approximated by

$$\begin{aligned}
Q_k &= \int_0^{t+\delta} \exp^{A\Delta t} GQG^{\top} \exp^{A^{\top}\Delta t} d\tau \\
&= \int_t^{t+\delta} \left(\sum_{k=0}^{\infty} \frac{\Delta t}{k!} A^k\right) GQG^{\top} \left(\sum_{k=0}^{\infty} \frac{\Delta t}{k!} A^k\right) d\tau \\
&= \int_t^{t+\delta} \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2} + \cdots + \frac{A^{\infty}(\Delta t)^{\infty}}{\infty!}\right) GQG^{\top} \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2} + \cdots + \frac{A^{\infty}(\Delta t)^{\infty}}{\infty!}\right)^{\top} d\tau \\
&\approx \int_t^{t+\delta} \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2} + \frac{A^3(\Delta t)^3}{2}\right) GQG^{\top} \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2} + \frac{A^3(\Delta t)^3}{2}\right)^{\top} d\tau
\end{aligned}$$

where $\Delta t = t + \delta - \tau$. Removing higher order terms we get

$$Q_k \approx \int_t^{t+\delta} \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2} + \frac{A^3(\Delta t)^3}{2}\right) GQG^{\top} \left(I + A\Delta t + \frac{A^2(\Delta t)^2}{2} + \frac{A^3(\Delta t)^3}{2}\right)^{\top} d\tau$$

expanding the equation we get

$$
\begin{aligned}
Q_k \approx \int_t^{t+\delta} (& IHI + IHA^\top \Delta t + \frac{IHA^{\top^2}(\Delta t)^2}{2} + \frac{IHA^{\top^3}(\Delta t)^3}{2} \\
& + A\Delta t HI + A\Delta t HA^\top \Delta t + \frac{A\Delta t HA^{\top^2}(\Delta t)^2}{2} + \frac{A\Delta t HA^{\top^3}(\Delta t)^3}{6} \\
& + \frac{A^2(\Delta t)^2 HI}{2} + \frac{A^2(\Delta t)^2 HA^\top \Delta t}{2} + \frac{A^2(\Delta t)^2}{2} H \frac{A^{\top^2}(\Delta t)^2}{2} + \frac{A^2(\Delta t)^2}{2} H \frac{A^{\top^3}(\Delta t)^3}{6} \\
& + \frac{A^3(\Delta t)^3 HI}{6} + \frac{A^3(\Delta t)^3 HA^\top \tau}{6} + \frac{A^3(\Delta t)^3}{6} H \frac{A^{\top^2}(\Delta t)^2}{2} + \frac{A^3(\Delta t)^3}{6} H \frac{A^{\top^3}(\Delta t)^3}{6}) d\tau
\end{aligned}
$$

where $H = GQG^\top$ .Removing higher order terms we get

$$
\begin{aligned}
Q_k \approx \int_t^{t+\delta} (& H + HA^\top \Delta t + \frac{HA^{\top^2}(\Delta t)^2}{2} + \frac{HA^{\top^3}(\Delta t)^3}{2} \\
& + AH\Delta t + AHA^\top \Delta t^2 + \frac{AHA^{\top^2}(\Delta t)^3}{2} \\
& + \frac{A^2(\Delta t)^2 H}{2} + \frac{A^2 HA^\top \Delta t^3}{2} + \frac{A^3 H(\Delta t)^3}{6}) d\tau
\end{aligned}
$$

which is integrated

$$
\begin{aligned}
Q_k \approx H\delta & + \frac{HA^\top \delta^2}{2} + \frac{HA^{\top^2}(\delta)^3}{6} + \frac{HA^{\top^3}(\delta)^4}{8} \\
& + \frac{AH\delta^2}{2} + \frac{AHA^\top \delta^3}{3} + \frac{AHA^{\top^2}(\delta)^4}{8} \\
& + \frac{A^2(\delta)^3 H}{6} + \frac{A^2 HA^\top \delta^4}{8} + \frac{A^3 H(\delta)^4}{24}
\end{aligned}
$$

In summary, we made the assumptions that the input $u(t)$ and $\mathbf{w}(t)$ are constant over the time interval δ in order to convert the continuous system to a discrete system. The end result is a system

$$
\begin{aligned}
x_k &= A_k x_{k-1} + B_k u_k + G\mathbf{w}_k && (24) \\
y_k &= C_k x + \mathbf{v}_k && (25)
\end{aligned}
$$

with prediction

$$
\begin{aligned}
\hat{x}_{k|j} &= A_k \hat{x}_{k-1|j} + B_k u_k \\
P_{k|j} &= A_k P_{k-1|j} A_k^\top + Q_k
\end{aligned}
$$

and update

$$
\begin{aligned}
K_k &= P_k C_k^\top \left(C_k P_k C_k^\top + R_k \right)^{-1} \\
\hat{x}_{k|k=j} &= \hat{x}_{k|j} + K_k \left(y_k - C_k \hat{x}_{k|j} \right) \\
P_{k|j} &= \left(I - K_k C_k \right) P_{k|j}
\end{aligned}
$$

where

$$A_k = \exp^{A\delta}$$

$$B_k = A^{-1}\left(\exp^{A\delta} - I\right) B$$

$$G_k = A^{-1}\left(\exp^{A\delta} - I\right) G$$

$$Q_k = \int_t^{t+\delta} \exp^{A(t+\delta-\tau)} GQG^\top \exp^{A^\top(t+\delta-\tau)} \, d\tau$$

which can be approximated as

$$A_k \approx I + A\delta + \frac{(A\delta)^2}{2} + \frac{(A\delta)^3}{6}$$

$$B_k \approx \left(2I + A\delta + \frac{A^2\delta^2}{3} + \frac{A^3\delta^3}{12}\right)\frac{\delta}{2}B$$

$$G_k \approx \left(2I + A\delta + \frac{A^2\delta^2}{3} + \frac{A^3\delta^3}{12}\right)\frac{\delta}{2}G$$

$$Q_k \approx H\delta + \frac{HA^\top\delta^2}{2} + \frac{HA^{\top^2}(\delta)^3}{6} + \frac{HA^{\top^3}(\delta)^4}{8} + \frac{AH\delta^2}{2} + \frac{AHA^\top\delta^3}{3} + \frac{A^2H(\delta)^3}{6} + \frac{AHA^{\top^2}(\delta)^4}{8} + \frac{A^2HA^\top\delta^4}{8} + \frac{A^3H(\delta)^4}{24}$$

Note that $Q_k = G_kQG_k^\top$

3.2 Note on Approximations

In some applications, the time step may vary or the system matrix A might not be invertible. In these cases, it is good to use the approximations when it is demanded by either A being singular or for computational needs. Also, in some systems, higher orders of A may be zero. In these cases, you can calculate the true values using approximation methods.

References

[1] L. Ljung, "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," IEEE Trans. Automat. Contr., AC-24, pp. 36–50, 1979.

[2] H. Kwakernaak and R. Sivan, Linea Optimal Control Systems, Wley-Interscience, New York, 1972.