Lecture 16: Extended Kalman Filter

The *extended Kalman filter (EKF)* is a state estimator for nonlinear systems that extends the Kalman filter for linear systems to nonlinear ones. Consider the following continuous-time nonlinear system:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{w}, t) \tag{1}$$
$$\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{v}, t) \tag{2}$$
$$\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_c) \tag{3}$$
$$\boldsymbol{v} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_c) \tag{4}$$

where, as usual, $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{u} \in \mathbb{R}^m$, and $\boldsymbol{y} \in \mathbb{R}^p$. The process noise $\boldsymbol{w} \in \mathbb{R}^w$ and measurement noise $\boldsymbol{v} \in \mathbb{R}^v$ are white, Gaussian, and zero-mean and have covariance matrices $\boldsymbol{Q} \in \mathbb{R}^{w \times w}$ and $\boldsymbol{R} \in \mathbb{R}^{v \times v}$. Notice, however, that the noise is not necessarily *additive*. The basic idea behind the EKF is straightforward — we will linearize the above system and then apply the conventional (linear) Kalman Filter. The resulting filter is no longer optimal but regardless of this fact it is still the de facto standard in many state estimation applications. The one question that remains is what trajectory to linearize our system around? The EKF uses a linearization around the current estimated state — in effect the system is "re-linearized" throughout the estimation procedure. Before delving into further details we first consider a simpler case wherein the linearization point is chosen before-hand.

Continuous-Time Linearized Kalman Filter (CT-LKF) [Simon, 2006, Ch.13]. Recall that we can linearize the nonlinear system (1)–(4) about a nominal (pre-determined) trajectory

$$(\boldsymbol{x}^{\text{ref}}(t), \boldsymbol{u}^{\text{ref}}(t), \boldsymbol{w}^{\text{ref}}(t), \boldsymbol{v}^{\text{ref}}(t), \boldsymbol{y}_{\text{ref}}(t)) \tag{5}$$

that satisfies

$$\dot{\boldsymbol{x}}^{\text{ref}} = \boldsymbol{f}(\boldsymbol{x}^{\text{ref}}, \boldsymbol{u}^{\text{ref}}, \boldsymbol{w}^{\text{ref}}, t) \tag{6}$$
$$\boldsymbol{y}^{\text{ref}} = \boldsymbol{h}(\boldsymbol{x}^{\text{ref}}, \boldsymbol{v}^{\text{ref}}, t) \tag{7}$$

by Taylor series expansion:

$$\dot{\boldsymbol{x}} \approx \boldsymbol{f}(\boldsymbol{x}^{\text{ref}}, \boldsymbol{u}^{\text{ref}}, \boldsymbol{w}^{\text{ref}}, t) + \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{f}\Big|_{\text{ref}}(\boldsymbol{x} - \boldsymbol{x}^{\text{ref}}) + \boldsymbol{J}_{\boldsymbol{u}}\boldsymbol{f}\Big|_{\text{ref}}(\boldsymbol{u} - \boldsymbol{u}^{\text{ref}}) + \boldsymbol{J}_{\boldsymbol{w}}\boldsymbol{f}\Big|_{\text{ref}}(\boldsymbol{w} - \boldsymbol{w}^{\text{ref}}) \tag{8}$$

$$\boldsymbol{y} \approx \boldsymbol{h}(\boldsymbol{x}^{\text{ref}}, \boldsymbol{v}^{\text{ref}}, t) + \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{f}\Big|_{\text{ref}}(\boldsymbol{x} - \boldsymbol{x}^{\text{ref}}) + \boldsymbol{J}_{\boldsymbol{v}}\boldsymbol{f}\Big|_{\text{ref}}(\boldsymbol{v} - \boldsymbol{v}^{\text{ref}}) \tag{9}$$

where the $\big|_*$ denotes evaluation along the trajectory (5). Introduce the deviation variables (our "tilde variables") as

$$\tilde{\boldsymbol{x}}(t) = \boldsymbol{x}(t) - \boldsymbol{x}^{\text{ref}}(t) \tag{10}$$
$$\tilde{\boldsymbol{u}}(t) = \boldsymbol{u}(t) - \boldsymbol{u}^{\text{ref}}(t) \tag{11}$$
$$\tilde{\boldsymbol{w}}(t) = \boldsymbol{w}(t) - \underbrace{\boldsymbol{w}^{\text{ref}}(t)}_{\text{typically zero}} \tag{12}$$
$$\tilde{\boldsymbol{y}}(t) = \boldsymbol{y}(t) - \underbrace{\boldsymbol{y}^{\text{ref}}(t)}_{=\boldsymbol{h}(\boldsymbol{x}^{\text{ref}},\boldsymbol{v}^{\text{ref}},t)} \tag{13}$$
$$\tilde{\boldsymbol{v}}(t) = \boldsymbol{v}(t) - \underbrace{\boldsymbol{v}^{\text{ref}}(t)}_{\text{typically zero}} \tag{14}$$

and denote each Jacobian as

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{J_x f}\Big|_{\boldsymbol{x}^*,\boldsymbol{u}^*,\boldsymbol{w}^{\text{ref}}} & \boldsymbol{B} &= \boldsymbol{J_u f}\Big|_{\boldsymbol{x}^*,\boldsymbol{u}^*,\boldsymbol{w}^{\text{ref}}} & \boldsymbol{L} &= \boldsymbol{J_w f}\Big|_{\boldsymbol{x}^*,\boldsymbol{u}^*,\boldsymbol{w}^{\text{ref}}} \\ \boldsymbol{C} &= \boldsymbol{J_x h}\Big|_{\boldsymbol{x}^*,\boldsymbol{u}^*,\boldsymbol{v}^{\text{ref}}} & \boldsymbol{M} &= \boldsymbol{J_v h}\Big|_{\boldsymbol{x}^*,\boldsymbol{u}^*,\boldsymbol{v}^{\text{ref}}} \end{aligned} \tag{15}$$

to obtain the linearized system

$$\dot{\tilde{\boldsymbol{x}}} = \boldsymbol{A}\tilde{\boldsymbol{x}} + \boldsymbol{B}\tilde{\boldsymbol{u}} + \boldsymbol{L}\tilde{\boldsymbol{w}} \tag{16}$$
$$\tilde{\boldsymbol{y}} = \boldsymbol{C}\tilde{\boldsymbol{x}} + \boldsymbol{M}\tilde{\boldsymbol{v}} \tag{17}$$
$$\tilde{\boldsymbol{w}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_c) \tag{18}$$
$$\tilde{\boldsymbol{v}} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_c) \tag{19}$$

With a linear system in hand we can apply the continuous Kalman-Bucy filter equations or we can discretize it further and apply the discrete-time Kalman filter equations. In many cases the nominal value of the noise terms is zero, i.e., $\boldsymbol{v}^{\text{ref}}(t) = \boldsymbol{w}^{\text{ref}}(t) = 0$, as indicated above. Also, recall that our Kalman-Bucy equations derived previously did not consider the $\boldsymbol{L}$ and $\boldsymbol{M}$ matrices that transform the noise variables. However, this is not an issue since we can always rewrite (16)–(19) in the form needed for the Kalman-Bucy equations

$$\dot{\tilde{\boldsymbol{x}}} = \boldsymbol{A}\tilde{\boldsymbol{x}} + \boldsymbol{B}\tilde{\boldsymbol{u}} + \tilde{\boldsymbol{w}} \tag{20}$$
$$\tilde{\boldsymbol{y}} = \boldsymbol{C}\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{v}} \tag{21}$$
$$\tilde{\boldsymbol{w}} \sim (\boldsymbol{0}, \tilde{\boldsymbol{Q}}_c) \tag{22}$$
$$\tilde{\boldsymbol{v}} \sim (\boldsymbol{0}, \tilde{\boldsymbol{R}}_c) \tag{23}$$

by modifying the noise covariance matrices as

$$\tilde{\boldsymbol{Q}}_c = \boldsymbol{L}\boldsymbol{Q}_c\boldsymbol{L}^{\text{T}} \tag{24}$$
$$\tilde{\boldsymbol{R}}_c = \boldsymbol{M}\boldsymbol{R}_c\boldsymbol{M}^{\text{T}} \tag{25}$$

(recall our lecture on linear transformation of Gaussian random vectors). This approach gives rise to the *continuous-time linearized Kalman filter (CT-LKF)*. Suppose the initial state estimate and state estimate covariance are:

$$\hat{\tilde{\boldsymbol{x}}}(0) = E[\tilde{\boldsymbol{x}}(0)] \tag{26}$$

$$\boldsymbol{P}(0) = E[(\tilde{\boldsymbol{x}}(0) - \hat{\tilde{\boldsymbol{x}}}(0))(\tilde{\boldsymbol{x}}(0) - \hat{\tilde{\boldsymbol{x}}}(0))^{\mathrm{T}}] \tag{27}$$

then the continuous-time Kalman filter equations for this linearized system are the Kalman-Bucy equations

$$\dot{\hat{\tilde{\boldsymbol{x}}}} = \boldsymbol{A}\hat{\tilde{\boldsymbol{x}}} + \boldsymbol{B}\tilde{\boldsymbol{u}} + \boldsymbol{K}(\tilde{\boldsymbol{y}} - \boldsymbol{C}\hat{\tilde{\boldsymbol{x}}}) \tag{28}$$

$$\dot{\boldsymbol{P}} = \boldsymbol{A}\boldsymbol{P} + \boldsymbol{P}\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{P}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{R}}_c^{-1}\boldsymbol{C}\boldsymbol{P}^{-1} + \tilde{\boldsymbol{Q}}_c \tag{29}$$

where in the first equation $\boldsymbol{K} = \boldsymbol{P}\boldsymbol{C}^{\mathrm{T}}\tilde{\boldsymbol{R}}_c^{-1}$, the second equation is the differential Ricatti equation (DRE), and the system matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ $\tilde{\boldsymbol{Q}}_c$, and $\tilde{\boldsymbol{R}}_c$ are related to the nonlinear system (1)–(4) through (35) and (24)–(25). To recover the actual state estimate we add the nominal trajectory to the deviation

$$\hat{\boldsymbol{x}} = \boldsymbol{x}^{\mathrm{ref}} + \hat{\tilde{\boldsymbol{x}}} \tag{30}$$

The linearized Kalman filter approach is well suited for cases where the system will always stay close to a nominal trajectory. This many occur, for example, if the nominal trajectory is a constant state/control pair corresponding to a stable equilibrium point. However this approach cannot be applied to cases where it is difficult to predict the nominal trajectory (and guarantee that it stays close). To handle this situation the extended Kalman filter (EKF) can be used, as described next.

Continuous-time Extended Kalman Filter (CT-EKF). The continuous-time EKF is very closely related to the continue-time linearized KF described in the previous section. Begin by differentiating the deviation state vector $\hat{\boldsymbol{x}} = \boldsymbol{x}^{\mathrm{ref}} + \hat{\tilde{\boldsymbol{x}}}$ to obtain

$$\dot{\hat{\boldsymbol{x}}} = \dot{\boldsymbol{x}}^{\mathrm{ref}} + \dot{\hat{\tilde{\boldsymbol{x}}}} \tag{31}$$

and substitute (6) and (28) to give

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{f}(\boldsymbol{x}^{\mathrm{ref}}, \boldsymbol{u}^{\mathrm{ref}}, \boldsymbol{w}^{\mathrm{ref}}, t) + \boldsymbol{A}\hat{\tilde{\boldsymbol{x}}} + \boldsymbol{B}\tilde{\boldsymbol{u}} + \boldsymbol{K}(\tilde{\boldsymbol{y}} - \boldsymbol{C}\hat{\tilde{\boldsymbol{x}}}) \tag{32}$$

$$= \boldsymbol{f}(\boldsymbol{x}^{\mathrm{ref}}, \boldsymbol{u}^{\mathrm{ref}}, \boldsymbol{w}^{\mathrm{ref}}, t) + \boldsymbol{A}(\hat{\boldsymbol{x}} - \boldsymbol{x}^{\mathrm{ref}}) + \boldsymbol{B}(\boldsymbol{u} - \boldsymbol{u}^{\mathrm{ref}}) + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{y}^{\mathrm{ref}} - \boldsymbol{C}(\hat{\boldsymbol{x}} - \boldsymbol{x}^{\mathrm{ref}})) \tag{33}$$

Now, choosing the nominal/reference state to the to be the estimated state, $\boldsymbol{x}^{\mathrm{ref}}(t) = \hat{\boldsymbol{x}}(t)$ and choose the reference control to be the actual control $\boldsymbol{u}^{\mathrm{ref}}(t) = \boldsymbol{u}(t)$, and substitute $\boldsymbol{y}^{\mathrm{ref}}(t) = \boldsymbol{h}(\hat{\boldsymbol{x}}(t), \boldsymbol{v}^{\mathrm{ref}}(t), t)$ from (7), the expression (33) simplifies as

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{f}(\hat{\boldsymbol{x}}^{\mathrm{ref}}, \boldsymbol{u}, \boldsymbol{w}^{\mathrm{ref}}, t) + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{h}(\hat{\boldsymbol{x}}, \boldsymbol{v}^{\mathrm{ref}}, t)) \tag{34}$$

which is in a form similar to the Kalman-Bucy equation (28) but $\boldsymbol{A}\hat{\tilde{\boldsymbol{x}}} + \boldsymbol{B}\tilde{\boldsymbol{u}}$ in (28) is replaced with the nonlinear $\boldsymbol{f}(\hat{\boldsymbol{x}}, \boldsymbol{u}^{\mathrm{ref}}, \boldsymbol{w}^{\mathrm{ref}}, t)$. Note also that this equation does not involve deviation variables . Since we chose to linearize around an on-line estimate we are, in effect, repeatedly adjusting the linearization at each time step to more closely approximate the dynamics around the current state. This is, of course, more accurate compared to linearizing about a single point throughout the entire trajectory. The EKF linearized system matrices are therefore

$$\begin{aligned} \bar{\boldsymbol{A}}(\hat{\boldsymbol{x}}) &= \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{f}\Big|_{\hat{\boldsymbol{x}},\boldsymbol{u},\boldsymbol{w}^{\mathrm{ref}}} & \qquad \bar{\boldsymbol{L}}(\hat{\boldsymbol{x}}) &= \boldsymbol{J}_{\boldsymbol{w}}\boldsymbol{f}\Big|_{\hat{\boldsymbol{x}},\boldsymbol{u},\boldsymbol{w}^{\mathrm{ref}}} \\ \bar{\boldsymbol{C}}(\hat{\boldsymbol{x}}) &= \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{h}\Big|_{\hat{\boldsymbol{x}},\boldsymbol{u},\boldsymbol{v}^{\mathrm{ref}}} & \qquad \bar{\boldsymbol{M}}(\hat{\boldsymbol{x}}) &= \boldsymbol{J}_{\boldsymbol{v}}\boldsymbol{h}\Big|_{\hat{\boldsymbol{x}},\boldsymbol{u},\boldsymbol{v}^{\mathrm{ref}}} \end{aligned} \tag{35}$$

where we've introduced the overbar notation to distinguish these matrices from the ones used in the CTLKF. In particular, these matrices are re-evaluated with each iteration of the EKF and so we make this more explicit by writing them as a function of $\hat{\boldsymbol{x}}$. Then the CT-EKF equations are then obtained much like the linear KF (28)–(29):

$$\dot{\hat{\boldsymbol{x}}} = \boldsymbol{f}(\hat{\boldsymbol{x}}, \boldsymbol{u}, \boldsymbol{w}^{\text{ref}}, t) + \boldsymbol{K}(\boldsymbol{y} - \boldsymbol{h}(\hat{\hat{\boldsymbol{x}}}, v_0, t)) \tag{36}$$

$$\dot{\boldsymbol{P}} = \bar{\boldsymbol{A}}(\hat{\boldsymbol{x}})\boldsymbol{P} + \boldsymbol{P}\bar{\boldsymbol{A}}(\hat{\boldsymbol{x}})^{\text{T}} - \boldsymbol{P}\bar{\boldsymbol{C}}(\hat{\boldsymbol{x}})^{\text{T}}\tilde{\boldsymbol{R}}^{-1}\bar{\boldsymbol{C}}(\hat{\boldsymbol{x}})\boldsymbol{P}^{-1} + \tilde{\boldsymbol{Q}} \tag{37}$$

where the Kalman gain is $\boldsymbol{K} = \boldsymbol{P}\boldsymbol{C}^{\text{T}}\tilde{\boldsymbol{R}}^{-1}$ in the first equation.

Discrete-time Extended Kalman Filter (DT-EKF). All of the steps described above can be repeated starting from a nonlinear discrete-time system model:

$$\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{w}_{k-1}) \tag{38}$$

$$\boldsymbol{y}_k = \boldsymbol{h}_k(\boldsymbol{x}_k, \boldsymbol{v}_k) \tag{39}$$

$$\boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{Q}) \tag{40}$$

$$\boldsymbol{v}_k \sim \mathcal{N}(0, \boldsymbol{R}) \tag{41}$$

Rather than using the continuous-time Kalman-Bucy equations we use the discrete-time KF equations to compute the state estimate. At the kth time-step we wish to linearize the above system around the *previous best estimate* of the system dynamics (i.e., we construct a Taylor series approximation around the posterior $\hat{\boldsymbol{x}}_{k-1|k-1}$ from the $(k-1)$th time-step). As before, we assume that the input to system $\boldsymbol{u}_{k-1}$ is known. The input modifies the dynamics at each step thus we can think of $\boldsymbol{u}_{k-1}$ as a known time-varying known parameter in $\boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{w}_{k-1})$. Also assume that the process noise is zero $\boldsymbol{w}_{k-1} = 0$. Thus, the Taylor series expansion is:

$$\boldsymbol{x}_k \approx \boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, 0) + \boldsymbol{F}_{k-1}(\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1|k-1}) + \boldsymbol{L}_{k-1}\boldsymbol{w}_{k-1} \tag{42}$$

where

$$\boldsymbol{F}_{k-1} = \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{f}\Big|_{\hat{\boldsymbol{x}}_{k-1|k-1}} \tag{43}$$

$$\boldsymbol{L}_{k-1} = \boldsymbol{J}_{\boldsymbol{w}}\boldsymbol{f}\Big|_{\hat{\boldsymbol{x}}_{k-1|k-1}} \tag{44}$$

We can re-write this system as:

$$\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, 0) + \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} - \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{L}_{k-1}\boldsymbol{w}_{k-1} \tag{45}$$

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \underbrace{[\boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, 0) - \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1}]}_{\tilde{\boldsymbol{u}}_{k-1}} + \underbrace{\boldsymbol{L}_{k-1}\boldsymbol{w}_{k-1}}_{\tilde{\boldsymbol{w}}_{k-1}} \tag{46}$$

$$= \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \tilde{\boldsymbol{u}}_{k-1} + \tilde{\boldsymbol{w}}_{k-1} \tag{47}$$

where

$$\tilde{\boldsymbol{u}}_{k-1} = \boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, 0) - \boldsymbol{F}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1} \tag{48}$$

is a new pseudo control input and

$$\tilde{\boldsymbol{w}}_{k-1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{L}_k\boldsymbol{Q}\boldsymbol{L}_k^{\text{T}}) \tag{49}$$

is a new pseudo noise for the modified system (47). We can now repeat this process for the measurement equation. Expanding as a Taylor series:

$$\boldsymbol{y}_k \approx \boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k-1}, 0) + \boldsymbol{H}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k-1}) + \boldsymbol{M}_k \boldsymbol{v}_k \tag{50}$$

where

$$\boldsymbol{H}_k = \boldsymbol{J}_{\boldsymbol{x}} \boldsymbol{h}\Big|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{51}$$

$$\boldsymbol{M}_k = \boldsymbol{J}_{\boldsymbol{v}} \boldsymbol{h}\Big|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{52}$$

so that the approximation above further simplifies as

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{z}_k + \tilde{\boldsymbol{v}}_k \tag{53}$$

where

$$\tilde{\boldsymbol{z}}_k = [\boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k-1}, 0) - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k-1}] \tag{54}$$

is a known signal input and

$$\tilde{\boldsymbol{v}}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{M}_k \boldsymbol{R} \boldsymbol{M}_k^{\mathrm{T}}) \tag{55}$$

is a new pseudo noise for the modified system (47). To summarize, the nonlinear system has been rewritten as:

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1} \boldsymbol{x}_{k-1} + \tilde{\boldsymbol{u}}_{k-1} + \tilde{\boldsymbol{w}}_{k-1} \tag{56}$$

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{z}_k + \tilde{\boldsymbol{v}}_k \tag{57}$$

with

$$E[\tilde{\boldsymbol{w}}_k \tilde{\boldsymbol{w}}_k^T] = \boldsymbol{L}_k \boldsymbol{Q} \boldsymbol{L}_k^{\mathrm{T}} \tag{58}$$

$$E[\tilde{\boldsymbol{v}}_k \tilde{\boldsymbol{v}}_k^T] = \boldsymbol{M}_k \boldsymbol{R} \boldsymbol{M}_k^{\mathrm{T}} \tag{59}$$

where

$$\tilde{\boldsymbol{u}}_{k-1} = \boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, 0) - \boldsymbol{F}_{k-1} \hat{\boldsymbol{x}}_{k-1|k-1} \tag{60}$$

$$\tilde{\boldsymbol{z}}_k = [\boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k-1}, 0) - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k|k-1}] \tag{61}$$

are a pseudo control input and a known signal, respectively. This expression is almost exactly in the form assumed by the discrete-time Kalman filter (DT-KF) with the exception of the extra $\boldsymbol{z}_k$ term. However, since $\boldsymbol{z}_k$ is a known/deterministic signal it does not affect the covariance propagation of the Kalman filter.

Recall that in the DT-KF the motion update is performed followed by the measurement update. Because of the choice of pseudo-control input the state prediction with the seemingly linear system is equivalent to a prediction with the nonlinear system:

$$\text{State Prediction :} \quad \hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}_{k-1} \hat{\boldsymbol{x}}_{k-1|k-1} + \tilde{\boldsymbol{u}}_{k-1} \tag{62}$$

$$= \boldsymbol{F}_{k-1} \hat{\boldsymbol{x}}_{k-1|k-1} + [\boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, 0) - \boldsymbol{F}_{k-1} \hat{\boldsymbol{x}}_{k-1|k-1}] \tag{63}$$

$$= \boldsymbol{f}_{k-1}(\hat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{0}) \tag{64}$$

whereas the covariance prediction, which does not depend on the pseudo control input, uses only the linearized dynamics:

$$\text{Covariance Prediction :} \qquad \boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{L}_{k-1}\boldsymbol{Q}\boldsymbol{L}_{k-1}^{\mathrm{T}} \tag{65}$$

Similarly, for the measurement equation the newly introduced signal z_k is known exactly since it depends only on $\hat{\boldsymbol{x}}_{k|k-1}$ which was computed in the previous step. The measurement prediction with the seemingly-linear system also reduces to a prediction with the nonlinear measurement equation upon substiting in $\boldsymbol{z}_k$:

$$\text{Measurement Prediction :} \qquad \boldsymbol{y}_{k|k-1} = \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{z}_k + \tilde{\boldsymbol{v}}_k \tag{66}$$

$$= \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1} + [\boldsymbol{h}_k(\hat{\boldsymbol{x}}_{k|k-1}, 0) - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k-1}] \tag{67}$$

$$= \boldsymbol{h}_k(\boldsymbol{x}_{k|k-1}, \boldsymbol{0}) \tag{68}$$

The measurement update step is then performed for the system linearizing around the the *predicted state* $\hat{\boldsymbol{x}}_{k|k-1}$ in (51) and (52). Thus, the innovation covariance and Kalman gain are:

$$\text{Innovation Covariance :} \qquad \boldsymbol{S}_k = \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{M}_k\boldsymbol{R}_k\boldsymbol{M}_k^T \tag{69}$$

$$\text{Kalman Gain :} \qquad \boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{S}_k^{-1} \tag{70}$$

Now when computing the posterior

$$\text{State Posterior :} \qquad \hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{y}_{k|k-1}) \tag{71}$$

$$\text{Covariance Posterior :} \qquad \boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1} \tag{72}$$

The known input signal $\boldsymbol{z}_k$ does not influence the uncertainty propagation when computing the covariance posterior since it is a known/deterministic vector.

Algorithm: Discrete-time Extended Kalman Filter

1. Ensure your system is in the form:

$$\boldsymbol{x}_k = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{w}_{k-1}) \tag{73}$$

$$\boldsymbol{y}_k = \boldsymbol{h}_k(\boldsymbol{x}_k, \boldsymbol{v}_k) \tag{74}$$

$$\boldsymbol{w}_k \sim \mathcal{N}(0, \boldsymbol{Q}) \tag{75}$$

$$\boldsymbol{v}_k \sim \mathcal{N}(0, \boldsymbol{R}) \tag{76}$$

where the system dynamics $\boldsymbol{f}_k$, measurement equation $\boldsymbol{h}_k$, and control input $\boldsymbol{u}_k$ is known for all $k = 0, 1, 2, \ldots$. The process and measurement noise covariances $\boldsymbol{Q}$ and $\boldsymbol{R}$ are also known where

$$E[\boldsymbol{w}_k\boldsymbol{w}_k^T] = \boldsymbol{Q} \tag{77}$$

$$E[\boldsymbol{v}_k\boldsymbol{v}_k^T] = \boldsymbol{R} \tag{78}$$

and it is assumed that $E[\boldsymbol{w}_k] = 0$ and $E[\boldsymbol{v}_k] = 0$ for all $k = 0, 1, 2, \ldots$ and $E[\boldsymbol{w}_k\boldsymbol{w}_q^T] = 0$ and $E[\boldsymbol{v}_k\boldsymbol{v}_q^T] = 0$ for all $q \neq k$.

2. Select an initial guess

$$\hat{\boldsymbol{x}}_{0|0} = E[\boldsymbol{x}_0] \tag{79}$$

$$\boldsymbol{P}_{0|0} = E[(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_{0|0})(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_{0|0})^{\mathrm{T}}] \tag{80}$$

for the filter and an initial covariance. In the absence of other information you may choose $\hat{\boldsymbol{x}}_0 = \boldsymbol{0}$ and $\boldsymbol{P}_0 = \boldsymbol{1}_{n\times n}\sigma_0^2$ where σ_0^2 is a large number (e.g., 10E6).

3. For $k = 1, 2, 3, \ldots$ perform the following:

   (a) Obtain the measurement $\boldsymbol{y}_k$ from the sensor/data stream.

   (b) Retrieve the posterior from the last step — this becomes the current state and covariance prior $\hat{\boldsymbol{x}}_{k-1|k-1}$ and $\boldsymbol{P}_{k-1|k-1}$.

   (c) Re-linearize the motion-update equations by computing:

$$\boldsymbol{F}_{k-1} = \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{f}\Big|_{\hat{\boldsymbol{x}}_{k-1|k-1}} \tag{81}$$

$$\boldsymbol{L}_{k-1} = \boldsymbol{J}_{\boldsymbol{w}}\boldsymbol{f}\Big|_{\hat{\boldsymbol{x}}_{k-1|k-1}} \tag{82}$$

   (d) Motion update: compute the following

$$\text{State Prediction :} \quad \hat{\boldsymbol{x}}_{k|k-1} = \boldsymbol{f}_{k-1}(\boldsymbol{x}_{k-1|k-1}, \boldsymbol{u}_{k-1}, \boldsymbol{0}) \tag{83}$$

$$\text{Covariance Prediction :} \quad \boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{L}_{k-1}\boldsymbol{Q}\boldsymbol{L}_{k-1}^{\mathrm{T}} \tag{84}$$

   (e) Re-linearize the measurement-update equations by computing:

$$\boldsymbol{H}_k = \boldsymbol{J}_{\boldsymbol{x}}\boldsymbol{h}\Big|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{85}$$

$$\boldsymbol{M}_k = \boldsymbol{J}_{\boldsymbol{v}}\boldsymbol{h}\Big|_{\hat{\boldsymbol{x}}_{k|k-1}} \tag{86}$$

   (f) Measurement update: compute the following

$$\text{Measurement Prediction :} \quad \boldsymbol{y}_{k|k-1} = \boldsymbol{h}_k(\boldsymbol{x}_{k|k-1}, \boldsymbol{0}) \tag{87}$$

$$\text{Innovation Covariance :} \quad \boldsymbol{S}_k = \boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{M}_k\boldsymbol{R}_k\boldsymbol{M}_k^T \tag{88}$$

$$\text{Kalman Gain :} \quad \boldsymbol{K}_k = \boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T\boldsymbol{S}_k^{-1} \tag{89}$$

$$\text{State Posterior :} \quad \hat{\boldsymbol{x}}_{k|k} = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{y}_{k|k-1}) \tag{90}$$

$$\text{Covariance Posterior :} \quad \boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H}_k)\boldsymbol{P}_{k|k-1} \tag{91}$$

   (g) Store the posterior and covariance to be used as the prior in the next iteration.

References

[Simon, 2006] Simon, D. (2006). *Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches*. John Wiley & Sons.