APPLIED KALMAN FILTERING

Extended Kalman Filter

Dr. Robert H. Bishop, P.E.

UNIVERSITY of SOUTH FLORIDA

- The continuous-discrete extended Kalman filter (EKF) is a direct extension of the optimal linear Kalman filter and applicable to systems described by nonlinear system dynamics and sensor mathematical models
- It is a recursive data processing algorithm that is capable of asynchronous fusion of measurements from various sensors with potentially time-varying noise characteristics utilizing prior knowledge regarding the estimated state of the system and the associated error statistics
- The EKF is model-dependent, thus making accurate modeling of the sensors and the underlying dynamics key factors
- In order to accommodate the nonlinearities of the system, the EKF utilizes Taylor Series expansions of the nonlinear system dynamics and measurement models along the current estimated trajectory
- With proper tuning, the EKF can provide accurate state estimates
- Tuning of the EKF is an art form that utilizes our intuition regarding the trade-offs between process noise and measurement noise strengths supported by *monte carlo* analysis, error budgets, and sensitivity analysis

- The nonlinear system model is assumed to be of the form

$$\boxed{\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{w}(t)} \tag{1}$$

where $\mathbf{x}(t) \in \Re^n$ is the state vector, $\mathbf{u}(t) \in \Re^p$ is a vector of known external inputs, $\mathbf{w}(t) \in \Re^n$ is the process noise, $\mathbf{f}(\mathbf{x}(t)) \in \Re^n$ is a vector of time-varying nonlinear smooth functions representing the system dynamics, and $\mathbf{G}(t) \in \Re^{n\times p}$ is a matrix of time-varying smooth functions is the mapping between the known inputs and the state

- The process noise is assumed to be a white noise process with

$$\mathrm{E}\left\{\mathbf{w}(t)\right\} = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}\left\{\mathbf{w}(t)\mathbf{w}^{\mathsf{T}}(\tau)\right\} = \mathbf{Q}_s(t)\delta(t-\tau),$$

where $\mathbf{Q}_s(t) = \mathbf{Q}_s^{\mathsf{T}}(t) \geq \mathbf{0} \in \mathbb{R}^{n\times n}$ is given

UNIVERSITY of SOUTH FLORIDA

- The nonlinear measurement model is assumed to be of the form

$$\boxed{\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k} \tag{2}$$

where the subscript "k" denotes a discrete time measurement at time $t = t_k$, $\mathbf{h}_k(\mathbf{x}_k) \in \mathbb{R}^m$ is the nonlinear measurement model evaluated at the state $\mathbf{x}_k = \mathbf{x}(t_k)$, and $\mathbf{v}_k \in \mathbb{R}^m$ is the measurement noise

- The measurement noise is assumed to be a white noise sequence with

$$\mathrm{E}\left\{\mathbf{v}_k\right\} = \mathbf{0} \qquad \text{and} \qquad \mathrm{E}\left\{\mathbf{v}_k \mathbf{v}_j^T\right\} = \mathbf{R}_k \delta_{kj},$$

where $\mathbf{R}_k = \mathbf{R}_k^T > \mathbf{0} \in \mathbb{R}^{m \times m}$ is given

- It is assumed that the process noise and the measurement noise are uncorrelated

$$\mathrm{E}\left\{\mathbf{w}(t)\mathbf{v}_k^T\right\} = \mathbf{0} \quad \forall \quad t, t_k$$

- We define the state estimate as

$$\boxed{\hat{\mathbf{x}}(t) := \mathrm{E}\left\{\mathbf{x}(t)\right\}}$$

- The state estimation error is computed via

$$\boxed{\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)}$$

and the state estimation error covariance is

$$\boxed{\mathbf{P}(t) = \mathrm{E}\left\{\mathbf{e}(t)\mathbf{e}^{\mathsf{T}}(t)\right\}} \tag{3}$$

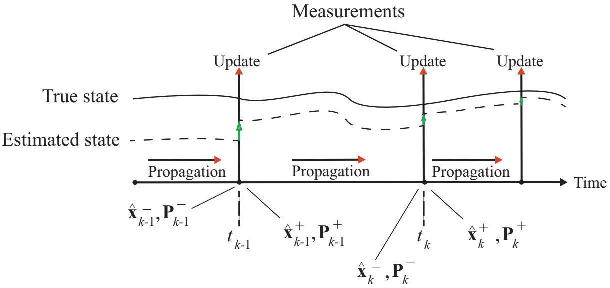
where we aim to design the EKF such that

$$\mathrm{E}\left\{\mathbf{e}(t)\right\} = \mathbf{0} \qquad \forall t \geq t_0$$

- Since this relationship holds for all time, it must be that the initial state estimate $\hat{\mathbf{x}}_0 = \mathrm{E}\left\{\mathbf{x}_0\right\}$ and we must choose the initial state estimation error covariance such that $\mathbf{P}_0 = \mathrm{E}\left\{\left(\mathbf{x}_0 - \hat{\mathbf{x}}_0\right)\left(\mathbf{x}_0 - \hat{\mathbf{x}}_0\right)^{\mathsf{T}}\right\}$

UNIVERSITY of SOUTH FLORIDA

- The EKF is a recursive algorithm comprised of two main phases: propagation and update.



- The state estimate and the state estimation error covariance are propagated forward to the time a discrete measurement becomes available at time t_k. At this time, the measurement is used to update both the state estimate and the state estimation error covariance
- The values immediately preceding an update are called as *a priori* and denoted with a superscript “–”, while the values immediately following an update are called as *a posteriori* and denoted with a superscript “+”

- Taking the expectation of each side in Eq. 1 yields

$$\mathrm{E}\left\{\dot{\mathbf{x}}(t)\right\} = \mathrm{E}\left\{\mathbf{f}(\mathbf{x}(t))\right\} + \mathbf{G}(t)\mathbf{u}(t) + \mathrm{E}\left\{\mathbf{w}(t)\right\}$$

- With $\dot{\hat{\mathbf{x}}}(t) = \mathrm{E}\left\{\dot{\mathbf{x}}(t)\right\}$ and expanding $\mathbf{f}(\mathbf{x}(t))$ in a Taylor Series expansion about $\hat{\mathbf{x}}(t)$ yields

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{F}(\hat{\mathbf{x}}(t))\mathrm{E}\left\{\mathbf{x}(t) - \hat{\mathbf{x}}(t)\right\} + \cdots + \mathbf{G}(t)\mathbf{u}(t) + \mathrm{E}\left\{\mathbf{w}(t)\right\}$$

where

$$\mathbf{F}(\hat{\mathbf{x}}(t)) = \left.\frac{\partial \mathbf{f}(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\right|_{\mathbf{x}(t)=\hat{\mathbf{x}}(t)}$$

- With $\mathrm{E}\left\{\mathbf{w}(t)\right\} = \mathbf{0}$, noting that we desire the EKF be an unbiased estimator $\mathrm{E}\left\{\mathbf{x}(t) - \hat{\mathbf{x}}(t)\right\} = 0$, and neglecting higher-order terms in the Taylor Series we have

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{G}(t)\mathbf{u}(t)} \tag{4}$$

- We are making the approximation that $\mathrm{E}\left\{\mathbf{f}(\mathbf{x}(t))\right\} = \mathbf{f}(\mathrm{E}\left\{\mathbf{x}(t)\right\})$ $\leftarrow$ engineering approximation

- From the definition of the estimation error in Eq. 3, we compute

$$\dot{\mathbf{e}}(t) = \dot{\mathbf{x}}(t) - \dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{w}(t) - \mathbf{f}(\hat{\mathbf{x}}(t))$$

- Utilizing a first-order Taylor Series expansion yields

$$\dot{\mathbf{e}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{F}(\hat{\mathbf{x}}(t))\,(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) + \cdots + \mathbf{w}(t) - \mathbf{f}(\hat{\mathbf{x}}(t))$$

- Neglecting the higher-order terms in the Taylor Series yields

$$\boxed{\dot{\mathbf{e}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\mathbf{e}(t) + \mathbf{w}(t)}$$

with the solution is

$$\boxed{\dot{\mathbf{e}}(t) = \Phi(t, t_0)\mathbf{e}(t_0) + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{w}(\tau)d\tau} \tag{5}$$

for all $t \geq t_0$, where the state transition matrix $\Phi(t, t_0)$ is the matrix associated with $\mathbf{F}(\hat{\mathbf{x}}(t))$ and satisfies the matrix differential equation

$$\dot{\Phi}(t, t_0) = \mathbf{F}(\hat{\mathbf{x}}(t))\Phi(t, t_0) \tag{6}$$

subject to the initial condition

$$\Phi(t_0, t_0) = \mathbf{I}$$

- Computing $\mathbf{P}(t) = E\{\mathbf{e}(t)\mathbf{e}^\top(t)\}$ in Eq. 3 using $\mathbf{e}(t)$ in Eq. 5 yields

$$\mathbf{P}(t) = E\{\mathbf{e}(t)\mathbf{e}^\top(t)\} = \Phi(t,t_0)E\{\mathbf{e}(t_0)\mathbf{e}^\top(t_0)\}\Phi^\top(t,t_0) + \int_{t_0}^{t}\int_{t_0}^{t}\Phi(t,\tau)E\{\mathbf{w}(\tau)\mathbf{w}^\top(\sigma)\}\Phi^\top(t,\sigma)d\sigma d\tau$$

  where we employ the assumption that $\mathbf{w}(t)$ is a white noise process uncorrelated with $\mathbf{e}(t_0)$

- With $\mathrm{E}\left\{\mathbf{w}(t)\mathbf{w}^\top(\tau)\right\} = \mathbf{Q}_s(t)\delta(t-\tau)$, $\forall t,\tau$, it follows that

$$\mathbf{P}(t) = \Phi(t,t_0)E\{\mathbf{e}(t_0)\mathbf{e}^\top(t_0)\}\Phi^\top(t,t_0) + \int_{t_0}^{t}\int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\delta(\tau-\sigma)\Phi^\top(t,\sigma)d\sigma d\tau$$

- Let $\mathbf{P}_0 = E\{\mathbf{e}(t_0)\mathbf{e}^\top(t_0)\}$ and by the property of the Dirac delta function we have

$$\boxed{\mathbf{P}(t) = \Phi(t,t_0)\mathbf{P}_0\Phi^\top(t,t_0) + \int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^\top(t,\tau)d\tau} \tag{7}$$

UNIVERSITY of SOUTH FLORIDA

- Recalling that $\dot{\Phi}(t,\tau) = \mathbf{F}(\hat{\mathbf{x}}(t))\Phi(t,\tau)$, $\mathbf{P}_0$ is constant and using Leibniz's Rule, we compute the time derivative of $\mathbf{P}(t)$ in Eq. 7 as

$$\dot{\mathbf{P}}(t) = \dot{\Phi}(t,t_0)\mathbf{P}_0\Phi^{\mathsf{T}}(t,t_0) + \Phi(t,t_0)\mathbf{P}_0\dot{\Phi}^{\mathsf{T}}(t,t_0) + \frac{d}{dt}\int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau$$

or

$$\dot{\mathbf{P}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\Phi(t,t_0)\mathbf{P}_0\Phi^{\mathsf{T}}(t,t_0) + \Phi(t,t_0)\mathbf{P}_0\Phi^{\mathsf{T}}(t,t_0)\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t)) + \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)\Big|_{\tau=t}$$
$$+\int_{t_0}^{t}\mathbf{F}(\hat{\mathbf{x}}(t))\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau + \int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t))d\tau$$

- Since $\mathbf{F}(\hat{\mathbf{x}}(t))$ is not a function of τ, we can take it outside the integration and collect terms yielding

$$\dot{\mathbf{P}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\left[\Phi(t,t_0)\mathbf{P}_0\Phi^{\mathsf{T}}(t,t_0) + \int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau\right] +$$
$$\left[\Phi(t,t_0)\mathbf{P}_0\Phi^{\mathsf{T}}(t,t_0) + \int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau\right]\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t)) + \mathbf{Q}_s(t)$$

- Substituting $\mathbf{P}(t)$ in Eq. 7 yields

$$\boxed{\dot{\mathbf{P}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t)) + \mathbf{Q}_s(t)} \qquad t \geq t_0 \tag{8}$$

  with the initial condition $\mathbf{P}_0 = \mathbf{P}_0^{\mathsf{T}} > 0$ given at $t = t_0$ and $\mathbf{Q}_s(t) = \mathbf{Q}_s^{\mathsf{T}}(t) \geq 0$ given
- Since $\mathbf{P}_0 = \mathbf{P}_0^{\mathsf{T}} > 0$ and $\mathbf{Q}_s(t) = \mathbf{Q}_s^{\mathsf{T}}(t) \geq 0$, it follows that $\mathbf{P}(t) = \mathbf{P}^{\mathsf{T}}(t) > 0, \forall t \geq t_0$
- Note that $\mathbf{P}(t)$ in Eq. 8 is an approximation to the true state estimation error covariance due to our use of Taylor Series approximations
- We will use various tools (e.g., *monte carlo*) to confirm the accuracy of the EKF state estimation error covariance

- Define the matrix $\bar{\mathbf{Q}}(t) = \bar{\mathbf{Q}}^{\mathsf{T}}(t) \geq \mathbf{0} \in \mathbb{R}^{n \times n}$ as

$$\bar{\mathbf{Q}}(t) = \int_{t_0}^{t} \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau$$

- Taking the time derivative of $\bar{\mathbf{Q}}(t)$ yields

$$\boxed{\dot{\bar{\mathbf{Q}}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\bar{\mathbf{Q}}(t) + \bar{\mathbf{Q}}(t)\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t)) + \mathbf{Q}_s(t)} \tag{9}$$

with the initial condition

$$\bar{\mathbf{Q}}(t_0) = \mathbf{0}$$

UNIVERSITY of SOUTH FLORIDA

- Consider the time interval $t \in \{t_{k-1}, t_k\} \rightarrow$ let $t_0 = t_{k-1}$
- We propagate the state estimate, $\hat{\mathbf{x}}(t)$, in Eq. 4, the state transition matrix, $\Phi(t, t_0)$, in Eq. 6, and the process noise, $\bar{\mathbf{Q}}(t)$, in Eq. 9 via

$$\boxed{\begin{aligned} \dot{\hat{\mathbf{x}}}(t) &= \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{G}(t)\mathbf{u}(t) \\ \dot{\Phi}(t, t_{k-1}) &= \mathbf{F}(\hat{\mathbf{x}}(t))\Phi(t, t_{k-1}) \\ \dot{\bar{\mathbf{Q}}}(t) &= \mathbf{F}(\hat{\mathbf{x}}(t))\bar{\mathbf{Q}}(t) + \bar{\mathbf{Q}}(t)\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t)) + \mathbf{Q}_s(t) \end{aligned}} \tag{10}$$

  with initial conditions $\hat{\mathbf{x}}(t_{k-1}) = \hat{\mathbf{x}}_{k-1}^{+}$, $\Phi(t_{k-1}, t_{k-1}) = \mathbf{I}$, and $\bar{\mathbf{Q}}(t_{k-1}) = \mathbf{0}$
- After integration, we set $\hat{\mathbf{x}}_k^{-} = \hat{\mathbf{x}}(t_k)$, $\Phi_{k-1} = \Phi(t_k, t_{k-1})$, and $\mathbf{Q}_{k-1} = \bar{\mathbf{Q}}(t_k)$ and using Eq. 7 we find

$$\boxed{\mathbf{P}_k^{-} = \Phi_{k-1}\mathbf{P}_{k-1}^{+}\Phi_{k-1}^{\mathsf{T}} + \mathbf{Q}_{k-1}} \tag{11}$$

  where $\mathbf{P}_k^{-} = \mathrm{E}\left\{\mathbf{e}_k^{-}(\mathbf{e}_k^{-})^{\mathsf{T}}\right\}$ and $\mathbf{P}_{k-1}^{+} = \mathrm{E}\left\{\mathbf{e}_{k-1}^{+}(\mathbf{e}_{k-1}^{+})^{\mathsf{T}}\right\}$

- Suppose that the state estimate and the state estimation error covariance are propagated forward to the time a discrete measurement becomes available
- At this time, the measurement is used to update both the state estimate and the state estimation error covariance
- Recall the nonlinear measurement model from Eq. 2 given by

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k$$

  Taking the expectation yields

$$\hat{\mathbf{y}}_k = \mathrm{E}\left\{\mathbf{y}_k\right\} = \mathrm{E}\left\{\mathbf{h}_k(\mathbf{x}_k)\right\} + \mathrm{E}\left\{\mathbf{v}_k\right\}$$

- Using a Taylor Series expansion of $\mathbf{h}_k(\mathbf{x}_k)$ about the *a priori* state estimate, $\hat{\mathbf{x}}_k^-$, yields

$$\hat{\mathbf{y}}_k = \mathrm{E}\left\{\mathbf{h}_k(\hat{\mathbf{x}}_k^-) + \mathbf{H}_k(\hat{\mathbf{x}}_k^-)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right) + \cdots\right\} + \mathrm{E}\left\{\mathbf{v}_k\right\} \tag{12}$$

  where $\mathbf{H}_k(\hat{\mathbf{x}}_k^-)$ is the measurement sensitivity matrix, defined to be

$$\mathbf{H}_k(\hat{\mathbf{x}}_k^-) = \left.\frac{\partial \mathbf{h}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k^-}$$

- Neglecting the higher-order terms in the Taylor Series expansion and noting that the EKF is an unbiased estimator (i.e. $\mathrm{E}\left\{\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right)\right\} = \mathbf{0}$) and that the measurement noise is assumed to be zero-mean (i.e. $\mathrm{E}\left\{\mathbf{v}_k\right\} = \mathbf{0}$), the estimated measurement is found to be

$$\boxed{\hat{\mathbf{y}}_k = \mathbf{h}_k(\hat{\mathbf{x}}_k^-)} \tag{13}$$

- We make the approximation that $\mathrm{E}\left\{\mathbf{h}(\mathbf{x}_k)\right\} = \mathbf{h}(\mathrm{E}\left\{\mathbf{x}_k\right\}) = \mathbf{h}(\hat{\mathbf{x}}_k^-)$ $\leftarrow$ engineering approximation
- Note that both the measurement and the estimated measurement are evaluated just prior to the measurement update

UNIVERSITY of SOUTH FLORIDA

- As with the Kalman filter development, begin with the Principle of Statistical Weighting. However, consider a variation to account for the fact that we are encountering a nonlinear setting

$$\boxed{\hat{\mathbf{x}}_k^+ = \mathbf{a}_k(\hat{\mathbf{x}}_k^-) + \mathbf{K}_k \mathbf{y}_k} \tag{14}$$

- How do we choose $\mathbf{a}_k(\hat{\mathbf{x}}_k^-)$ and $\mathbf{K}_k$ in this nonlinear setting?
- Define the estimation errors:

$$\boxed{\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-} \text{ (\textit{a priori}) and } \boxed{\mathbf{e}_k^+ = \mathbf{x}_k - \hat{\mathbf{x}}_k^+} \text{ (\textit{a posteriori})}$$

- We want to compute $\mathbf{a}_k(\hat{\mathbf{x}}_k^-)$ and $\mathbf{K}_k$ so that the EKF is an <u>unbiased</u> filter

$$E\{\mathbf{e}_k^-\} = 0 \quad \text{and} \quad E\{\mathbf{e}_k^+\} = 0 \quad \forall k \quad \text{and} \quad \forall \mathbf{x}_k$$

UNIVERSITY of SOUTH FLORIDA

- Start with Eq. 14 and compute

$$\mathbf{x}_k - \hat{\mathbf{x}}_k^+ = \mathbf{x}_k - \mathbf{a}_k(\hat{\mathbf{x}}_k^-) - \mathbf{K}_k \mathbf{y}_k$$

and using the definitions of $\mathbf{e}_k^+$, $\mathbf{e}_k^-$, and the measurement $\mathbf{y}_k$ in Eq. 13 it follows that

$$\mathbf{e}_k^+ = \mathbf{e}_k^- + \hat{\mathbf{x}}_k^- - \mathbf{a}_k(\hat{\mathbf{x}}_k^-) - \mathbf{K}_k \mathbf{h}(\mathbf{x}_k) - \mathbf{K}_k \mathbf{v}_k \tag{15}$$

- Taking the expectation yields

$$E\{\mathbf{e}_k^+\} = E\{\mathbf{e}_k^-\} + \hat{\mathbf{x}}_k^- - \mathbf{a}_k(\hat{\mathbf{x}}_k^-) - \mathbf{K}_k E\{\mathbf{h}(\mathbf{x}_k)\} - \mathbf{K}_k E\{\mathbf{v}_k\}\}$$

- We assume that $E\{\mathbf{v}_k\} = \mathbf{0}$ and $E\{\mathbf{e}_k^-\} = \mathbf{0}$
- Since we want $E\{\mathbf{e}_k^+\} = \mathbf{0}$ for all $\mathbf{x}_k$, we require

$$\boxed{\mathbf{a}_k(\hat{\mathbf{x}}_k^-) = \hat{\mathbf{x}}_k^- - \mathbf{K}_k E\{\mathbf{h}(\mathbf{x}_k)\}}$$

- Expanding $\mathbf{h}(\mathbf{x}_k)$ in a Taylor Series about $\hat{\mathbf{x}}_k^-$ yields

$$\mathbf{a}_k(\hat{\mathbf{x}}_k^-) = \hat{\mathbf{x}}_k^- - \mathbf{K}_k \mathrm{E}\left\{\mathbf{h}_k(\hat{\mathbf{x}}_k^-) + \mathbf{H}_k(\hat{\mathbf{x}}_k^-)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right) + \cdots\right\}$$

UNIVERSITY of SOUTH FLORIDA

- Neglecting higher-order terms in the Taylor Series yields and with $E\{\mathbf{e}_k^-\} = \mathbf{0}$ it follows that

$$\mathbf{a}_k(\hat{\mathbf{x}}_k^-) = \hat{\mathbf{x}}_k^- - \mathbf{K}_k \mathbf{h}_k(\hat{\mathbf{x}}_k^-) \tag{16}$$

- Substituting $\mathbf{a}_k(\hat{\mathbf{x}}_k^-)$ in Eq. 16 into Eq. 14 yields the estimated state update

$$\boxed{\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)\right)} \tag{17}$$

- Note that the form of the EKF state estimate update in Eq. 17 is <u>similar</u> to the state estimate update for the Kalman filter

$$\boxed{\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \hat{\mathbf{x}}_k^-\right)}$$

- Keep in mind that the state estimate update in Eq. 17 is an approximation .... so we'll need to expertly apply the tuning process to make sure the state estimation error covariance represents reality

- Substituting $\mathbf{a}_k(\hat{\mathbf{x}}_k^-)$ in Eq. 16 into Eq. 14 yields

$$\mathbf{e}_k^+ = \mathbf{e}_k^- - \mathbf{K}_k \left(\mathbf{h}(\mathbf{x}_k) - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)\right) - \mathbf{K}_k \mathbf{v}_k$$

- Expanding $\mathbf{h}(\mathbf{x}_k)$ in a Taylor Series about $\hat{\mathbf{x}}_k^-$ and neglecting higher order-terms yields

$$\boxed{\mathbf{e}_k^+ = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k(\hat{\mathbf{x}}_k^-)\right] \mathbf{e}_k^- - \mathbf{K}_k \mathbf{v}_k} \tag{18}$$

- Define $\mathbf{P}_k^- = \mathrm{E}\left\{\mathbf{e}_k^-(\mathbf{e}_k^-)^\mathsf{T}\right\}$ and $\mathbf{P}_k^+ = \mathrm{E}\left\{\mathbf{e}_k^+(\mathbf{e}_k^+)^\mathsf{T}\right\}$. Computing $\mathbf{P}_k^+$ using Eq. 18 and taking the expectation yields

$$\boxed{\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k(\hat{\mathbf{x}}_k^-)\right] \mathbf{P}_k^- \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k(\hat{\mathbf{x}}_k^-)\right]^\mathsf{T} + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^\mathsf{T}} \tag{19}$$

where we use our assumption that $\mathbf{v}_k$ is a white noise sequence with $\mathrm{E}\left\{\mathbf{v}_k \mathbf{v}_k^\mathsf{T}\right\} = \mathbf{R}_k$ and $\mathrm{E}\left\{\mathbf{e}_k^- \mathbf{v}_k^\mathsf{T}\right\} = \mathbf{0}, \forall k$

- Notice that update $\mathbf{P}_k^+$ has the same form as the Joseph formula, but keep in mind that this is an approximation

UNIVERSITY of SOUTH FLORIDA

- As with the Kalman filter, our goal is to minimize the expected mean squared state estimation error at the update
- Define the performance index $\mathcal{J}_k = E\{(\mathbf{e}_k^+)^\top \mathbf{e}_k^+\}$. Then with $\mathbf{P}_k^+ := E\{\mathbf{e}_k^+ (\mathbf{e}_k^+)^\top\}$ it follows that

$$\boxed{\mathcal{J}_k = tr\left\{E\{\mathbf{e}_k^+ (\mathbf{e}_k^+)^\top\}\right\} = tr\ \mathbf{P}_k^+}$$

and

$$tr\ \mathbf{P}_k^+ = tr\ \mathbf{P}_k^- - 2tr\ \mathbf{K}_k\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^- + tr\ \mathbf{K}_k\left[\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^-\mathbf{H}_k^\top(\hat{\mathbf{x}}_k^-) + \mathbf{R}_k\right]\mathbf{K}_k^\top$$

- Our approach is to take the partial derivative of $tr\ \mathbf{P}_k^+$ with respect to $\mathbf{K}_k$, set the result to zero and solve for the EKF gain $\mathbf{K}_k$, thus

$$\frac{\partial}{\partial \mathbf{K}_k} tr\,\mathcal{J}_k = -2\mathbf{P}_k^-\mathbf{H}_k^\top(\hat{\mathbf{x}}_k^-) + 2\mathbf{K}_k\left[\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^-\mathbf{H}_k^\top(\hat{\mathbf{x}}_k^-) + \mathbf{R}_k\right] = \mathbf{0}$$

- Solving for the EKF gain matrix $\mathbf{K}_k$ yields

$$\boxed{\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^\top(\hat{\mathbf{x}}_k^-)\left[\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^-\mathbf{H}_k^\top(\hat{\mathbf{x}}_k^-) + \mathbf{R}_k\right]^{-1}} \qquad (20)$$

UNIVERSITY of SOUTH FLORIDA

- Substituting the EKF gain $\mathbf{K}_k$ in Eq. 20 into the the state estimation error covariance $\mathbf{P}_k^+$ in Eq. 19 yields

$$\boxed{\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k(\hat{\mathbf{x}}_k^-)\right]\mathbf{P}_k^-} \tag{21}$$

- Note that the form of the simplified EKF state estimation error covariance update in Eq. 21 is <u>similar</u> to the simplified state estimation error covariance update for the Kalman filter

$$\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right]\mathbf{P}_k^-$$

- While the EKF $\mathbf{P}_k^+$ has the same form as the simplified Kalman filter formula, keep in mind that this is an approximation

System model	$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{w}(t)$
Measurement model	$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k$
Noise model	$\mathrm{E}\left\{\mathbf{w}(t)\right\} = \mathbf{0},\ \mathrm{E}\left\{\mathbf{w}(t)\mathbf{w}^\mathsf{T}(\tau)\right\} = \mathbf{Q}_s(t)\delta(t-\tau)$ $\mathrm{E}\left\{\mathbf{v}_k\right\} = \mathbf{0},\ \mathrm{E}\left\{\mathbf{v}_k\mathbf{v}_j^\mathsf{T}\right\} = \mathbf{R}_k\delta_{kj}$ $\mathrm{E}\left\{\mathbf{w}(t)\mathbf{v}_k^\mathsf{T}\right\} = \mathbf{0},\ \forall t, k$
Initial conditions	$\mathbf{x}_0 = \mathbf{x}(t_0)$
Assumptions	$\mathbf{Q}_s(t) = \mathbf{Q}_s^\mathsf{T}(t) \geq \mathbf{0},\ \forall t$ and $\mathbf{R}_k = \mathbf{R}_k^\mathsf{T} > \mathbf{0},\ \forall k$

- The model $\mathbf{f}(\mathbf{x}(t)) \in \Re^n$ and matrices $\mathbf{G}(t) \in \Re^{n\times p}$ and $\mathbf{Q}_s(t) \in \Re^{n\times n}$ are given $\forall t \geq t_0$
- The model $\mathbf{h}_k(\mathbf{x}_k) \in \Re^m$ and matrix $\mathbf{R}_k \in \Re^{m\times m}$ are given $\forall k$
- The exogenous input $\mathbf{u}(t) \in \Re^p$ is known $\forall t \geq t_0$

Initial conditions	$\hat{\mathbf{x}}_0 = \mathrm{E}\left\{\mathbf{x}(t_0)\right\} \;\; \mathbf{P}_0 = \mathrm{E}\left\{(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^\mathsf{T}\right\}$
Propagate	$\hat{\mathbf{x}}(t_{k-1}) = \hat{\mathbf{x}}_{k-1}^+,\ \mathbf{P}(t_{k-1}) = \mathbf{P}_{k-1}^+$ $\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{G}(t)\mathbf{u}(t)$ $\dot{\mathbf{P}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^\mathsf{T}(\hat{\mathbf{x}}(t)) + \mathbf{Q}_s(t)$ $\hat{\mathbf{x}}_k^- := \hat{\mathbf{x}}(t_k)\ ,\ \mathbf{P}_k^- := \mathbf{P}(t_k)$
Update	$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}(\hat{\mathbf{x}}_k^-)\left[\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}(\hat{\mathbf{x}}_k^-) + \mathbf{R}_k\right]^{-1}$ $\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)\right)$ $\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\right]\mathbf{P}_k^-$
Definitions	$\mathbf{F}(\hat{\mathbf{x}}(t)) := \left.\frac{\partial\mathbf{f}(\mathbf{x}(t))}{\partial\mathbf{x}(t)}\right\vert_{\mathbf{x}(t)=\hat{\mathbf{x}}(t)},\ \mathbf{H}_k(\hat{\mathbf{x}}_k^-) := \left.\frac{\partial\mathbf{h}_k(\mathbf{x}_k)}{\partial\mathbf{x}_k}\right\vert_{\mathbf{x}_k=\hat{\mathbf{x}}_k^-}$

Initial conditions	$\hat{\mathbf{x}}_0 = \mathrm{E}\left\{\mathbf{x}(t_0)\right\} \; \mathbf{P}_0 = \mathrm{E}\left\{(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^{\mathsf{T}}\right\}$
Propagate	$\hat{\mathbf{x}}(t_{k-1}) = \hat{\mathbf{x}}^+_{k-1},\ \Phi(t_{k-1}) = \mathbf{I},\ \bar{\mathbf{Q}}(t_{k-1}) = \mathbf{0}$ $\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t)) + \mathbf{G}(t)\mathbf{u}(t)$ $\dot{\Phi}(t, t_{k-1}) = \mathbf{F}(\hat{\mathbf{x}}(t))\Phi(t, t_{k-1})$ $\dot{\bar{\mathbf{Q}}}(t) = \mathbf{F}(\hat{\mathbf{x}}(t))\bar{\mathbf{Q}}(t) + \bar{\mathbf{Q}}(t)\mathbf{F}^{\mathsf{T}}(\hat{\mathbf{x}}(t)) + \mathbf{Q}_s(t)$ $\Phi_{k-1} := \Phi(t_k, t_{k-1})\ ,\ \mathbf{Q}_{k-1} := \bar{\mathbf{Q}}(t_k)$ $\mathbf{P}^-_k = \Phi_{k-1}\mathbf{P}^+_{k-1}\Phi^{\mathsf{T}}_{k-1} + \mathbf{Q}_{k-1}$
Update	$\mathbf{K}_k = \mathbf{P}^-_k\mathbf{H}^{\mathsf{T}}_k(\hat{\mathbf{x}}^-_k)\left[\mathbf{H}_k(\hat{\mathbf{x}}^-_k)\mathbf{P}^-_k\mathbf{H}^{\mathsf{T}}_k(\hat{\mathbf{x}}^-_k) + \mathbf{R}_k\right]^{-1}$ $\hat{\mathbf{x}}^+_k = \hat{\mathbf{x}}^-_k + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}^-_k)\right)$ $\mathbf{P}^+_k = \left[\mathbf{I} - \mathbf{K}_k\mathbf{H}_k(\hat{\mathbf{x}}^-_k)\right]\mathbf{P}^-_k$
Definitions	$\mathbf{F}(\hat{\mathbf{x}}(t)) := \left.\frac{\partial \mathbf{f}(\mathbf{x}(t))}{\partial \mathbf{x}(t)}\right\vert_{\mathbf{x}(t)=\hat{\mathbf{x}}(t)},\ \mathbf{H}_k(\hat{\mathbf{x}}^-_k) := \left.\frac{\partial \mathbf{h}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k}\right\vert_{\mathbf{x}_k=\hat{\mathbf{x}}^-_k}$

UNIVERSITY of SOUTH FLORIDA

END MODULE