APPLIED KALMAN FILTERING

State-Space Systems and Solutions

Dr. Robert H. Bishop, P.E.

UNIVERSITY of SOUTH FLORIDA

- The dynamics of linear, lumped parameter systems can be represented by the first-order vector-matrix differential equation

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

  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, $\mathbf{F}(t) \in \Re^{n \times n}$ is a matrix of time-varying 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
- $\mathbf{F}(t)$ and $\mathbf{G}(t)$ are known matrices
- The state vector is composed of any set of quantities sufficient to completely describe the unforced motion of the system
- The state vector $\mathbf{x}(t)$ is not a unique set of variables. Any other set $\mathbf{z}(t)$ related to $\mathbf{x}(t)$ by a nonsingular transformation, $\mathbf{T}(t) \in \Re^{n \times n}$, given by

$$\boxed{\mathbf{z}(t) = \mathbf{T}(t)\mathbf{x}(t)} \text{ where } \mathbf{T}^{-1}(t) \text{ exists } \forall t$$

  also are sufficient to completely describe the unforced motion of the system

- In many practical situations, the real-world evolves in continuous-time and is typically effectively modeled with a set of linear, time-varying, smooth vector functions derived from physical principles
- Our mathematical model of the real-world is thus a continuous-time dynamical representation in which we include models of systematic errors and random uncertainties
- In many practical situations, the external measurements are available at only discrete points in time and again are typically effectively modeled with a set of linear, time-varying, smooth vector functions derived from physical principles based on the sensor type
- Our mathematical model of the sensors is thus a set of discrete-time representations in which we include models of systematic errors and random uncertainties

- Consider the case where $\mathbf{x}(t_0) = \mathbf{x}_0$, $\mathbf{u}(t) = 0$, and $\mathbf{w}(t) = 0$. Then it follows that

$$\boxed{\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t)}$$

- Define a <u>state transition matrix</u>, denoted by $\Phi(t,\tau) \in \Re^{n\times n}$, $\forall t, \tau$, such that

$$\boxed{\dot{\Phi}(t,\tau) = \mathbf{F}(t)\Phi(t,\tau)} \tag{2}$$

where $\dot{\Phi}(t,\tau) := \partial\Phi(t,\tau)/\partial t$ and $\Phi(t,t) = \mathbf{I}$

- <u>Claim</u>:

$$\boxed{\mathbf{x}(t) = \Phi(t,t_0)\mathbf{x}_0} \quad \forall t \geq t_0 \tag{3}$$

- Taking the time-derivative of $\mathbf{x}(t)$ in Eq. 3 and using Eq. 2 yields

$$\dot{\mathbf{x}}(t) = \dot{\Phi}(t,t_0)\mathbf{x}_0 = \mathbf{F}(t)\Phi(t,t_0)\mathbf{x}_0 = \mathbf{F}(t)\mathbf{x}(t) \quad \checkmark$$

- Consider the case where $\mathbf{x}(0) = \mathbf{0}$. Then it follows that

$$\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{w}(t)$$

- <u>Claim</u>:

$$\boxed{\mathbf{x}(t) = \int_{t_0}^{t} \Phi(t,\tau)\Big(\mathbf{G}(\tau)\mathbf{u}(\tau) + \mathbf{w}(\tau)\Big) d\tau} \tag{4}$$

- Taking the time-derivative of $\mathbf{x}(t)$ in Eq. 4 yields

$$\begin{aligned}
\dot{\mathbf{x}}(t) &= \frac{d}{dt}\left(\int_{t_0}^{t} \Phi(t,\tau)\Big(\mathbf{G}(\tau)\mathbf{u}(\tau) + \mathbf{w}(\tau)\Big) d\tau\right) \\
&= \mathbf{F}(t)\int_{t_0}^{t} \Phi(t,\tau)\Big(\mathbf{G}(\tau)\mathbf{u}(\tau) + \mathbf{w}(\tau)\Big) d\tau + \Big[\Phi(t,\tau)\Big(\mathbf{G}(\tau)\mathbf{u}(\tau) + \mathbf{w}(\tau)\Big)\Big]_{\tau=t} \\
&= \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{w}(t) \quad \checkmark
\end{aligned}$$

- Recall Leibniz rule:

$$\frac{d}{dt}\left(\int_{t_0}^{t} \mathbf{f}(t,\tau) d\tau\right) = \mathbf{f}(t,\tau)\Big|_{\tau=t} + \int_{t_0}^{t} \frac{\partial \mathbf{f}(t,\tau)}{\partial t}\, d\tau$$

- Employing the Principle of Superposition which applies to linear, time-varying systems yields

$$\boxed{\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{u}(\tau)d\tau + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{w}(\tau)d\tau} \tag{5}$$

- When $\mathbf{F}$ is a constant matrix, the state transition matrix is given by

$$\Phi(t, t_0) = \Phi(t - t_0) = e^{\mathbf{F}(t-t_0)}$$

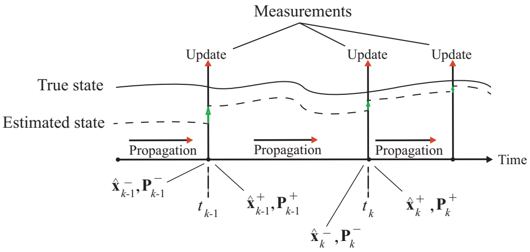
- When $\mathbf{F}(t)$ is time-varying, the state transition matrix

$$\Phi(t, t_0) \neq e^{\int_{t_0}^{t} \mathbf{F}(\tau)d\tau}$$

  except in the unlikely case that

$$\dot{\mathbf{F}}(t)\mathbf{F}(t) = \mathbf{F}(t)\dot{\mathbf{F}}(t)$$

- The Kalman filter has two main stages: propagation and update
- Propagation occurs between t_{k-1} and t_k, $\forall k \geq 1$ and the update occurs at t_k as measurements come available



- We need to consider how to use the mathematical models just described to create a discrete-time mathematical model including uncertainties (i.e., noise) that describes the state propagation between t_{k-1} and t_k, $\forall k$

- Consider the dynamical system described by the mathematical model in Eq. 1 where

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

  where $\mathbf{Q}_s(t) \in \Re^{n\times n}$ is a known matrix with $\mathbf{Q}_s(t) = \mathbf{Q}_s^{\mathsf{T}}(t) \geq \mathbf{0}$, $\forall t$
- Consider the discrete measurement at time t_k described by the mathematical model

$$\boxed{\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k},$$

  where $\mathbf{x}_k := \mathbf{x}(t_k)$, $\mathbf{H}_k := \mathbf{H}(t_k) \in \Re^{m\times n}$ is the measurement mapping matrix, and $\mathbf{v}_k \in \Re^{m}$ is the measurement noise with

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

  where $\mathbf{R}_k \in \Re^{m\times m}$ is the measurement noise covariance matrix with $\mathbf{R}_k = \mathbf{R}_k^{\mathsf{T}} > \mathbf{0}$, $\forall k$
- The known matrix $\mathbf{H}_k$ is comprised of time-varying smooth functions representing the mapping between the state and the measurement

- The solution of the dynamical equation creates the discrete-time mathematical model

$$\boxed{\mathbf{x}_k = \Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}},$$

where

$$\mathbf{u}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k,\tau)\mathbf{G}(\tau)\mathbf{u}(\tau)d\tau \in \Re^n, \quad \mathbf{w}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k,\tau)\mathbf{w}(\tau)d\tau \in \Re^n,$$

and $\Phi_{k-1} := \Phi(t_k, t_{k-1})$ is the state transition matrix with

$$\dot{\Phi}(t, t_{k-1}) = \mathbf{F}(t)\Phi(t, t_{k-1})\,, \quad t_{k-1} \leq t \leq t_k$$

where $\Phi(t_{k-1}, t_{k-1}) = \mathbf{I}$

- Note that

$$E\{\mathbf{w}_{k-1}\} = \int_{t_{k-1}}^{t_k} \Phi(t_k,\tau)E\{\mathbf{w}(\tau)\}d\tau = \mathbf{0}$$

$$\mathbf{Q}_{k-1} := E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^{\mathsf{T}}\} = \int_{t_{k-1}}^{t_k} \Phi(t_k,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t_k,\tau)d\tau$$

and $\mathbf{Q}_{k-1} = \mathbf{Q}_{k-1}^{\mathsf{T}} \geq \mathbf{0}$ (since $\mathbf{Q}_s(t) = \mathbf{Q}_s^{\mathsf{T}}(t) \geq \mathbf{0}$)

- Let's take a deeper look into $\boxed{\mathbf{Q}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\mathbf{Q}_s(\tau)\Phi^{\mathrm{T}}(t_k, \tau) d\tau}$
- Recall that we assumed $\mathbf{w}(t)$ is a zero-mean white noise process with

$$\mathrm{E}\left\{\mathbf{w}(t)\mathbf{w}^{\mathrm{T}}(\tau)\right\} = \mathbf{Q}_s(t)\delta(t-\tau), \qquad \forall t, \tau \tag{6}$$

- With $\mathbf{w}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\mathbf{w}(\tau)d\tau$, we have

$$\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\mathbf{w}(\tau)d\tau \int_{t_{k-1}}^{t_k} \mathbf{w}^T(\tau)\Phi^T(t_k, \tau)d\tau,$$

and

$$E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T\} = \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)E\{\mathbf{w}(\tau)\mathbf{w}^T(\sigma)\}\Phi^T(t_k, \sigma)d\sigma d\tau,$$

and using Eq. 6 it follows that

$$E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T\} = \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k} \Phi(t_k,\tau)\mathbf{Q}_s(\tau)\delta(\tau-\sigma)\Phi^T(t_k,\sigma)d\sigma d\tau,$$

- Employing the properties of the Dirac delta function $\delta(\tau-\sigma)$ in the integration over σ yields

$$\boxed{\mathbf{Q}_{k-1} := E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T\} = \int_{t_{k-1}}^{t_k} \Phi(t_k,\tau)\mathbf{Q}_s(\tau)\Phi^T(t_k,\tau)d\tau}$$

- Typically we do not explicitly compute $\mathbf{Q}_{k-1}$ using the integral above
- Let

$$\bar{\mathbf{Q}}(t) = \int_{t_{k-1}}^{t} \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau$$

- Consider Leibniz's Rule

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(t,\tau)d\tau = f(t,b(t))\frac{db(t)}{dt} - f(t,a(t))\frac{da(t)}{dt} + \int_{a(t)}^{b(t)} \frac{\partial f(t,\tau)}{\partial t}d\tau$$

- With $f(t,\tau) = \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)$, $a(t) = t_{k-1}$, and $b(t) = t$, we have

$$\dot{\bar{\mathbf{Q}}}(t) = \mathbf{Q}_s(t) + \int_{t_{k-1}}^{t} (\dot{\Phi}(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau) + \Phi(t,\tau)\mathbf{Q}_s(\tau)\dot{\Phi}^{\mathsf{T}}(t,\tau))d\tau$$

or

$$\dot{\bar{\mathbf{Q}}}(t) = \mathbf{Q}_s(t) + \int_{t_{k-1}}^{t} \left[\mathbf{F}(t)\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau) + \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)\mathbf{F}^{\mathsf{T}}(t)\right] d\tau$$

where we utilize the relationship

$$\dot{\Phi}(t,\tau) = \mathbf{F}(t)\Phi(t,\tau)\,, \quad \tau \le t$$

Re-arranging terms and noting that $\mathbf{F}(t)$ is not a function of τ hence can be taken outside the integral yields

$$\begin{aligned}\dot{\bar{\mathbf{Q}}}(t) \quad = \quad & \mathbf{Q}_s(t) + \mathbf{F}(t)\left[\int_{t_{k-1}}^{t} \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau\right] + \\ & \left[\int_{t_{k-1}}^{t} \Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau))d\tau\right]\mathbf{F}^{\mathsf{T}}(t)\end{aligned}$$

- It then follows from the definition of $\bar{\mathbf{Q}}(t)$ that

$$\boxed{\dot{\bar{\mathbf{Q}}}(t) = \mathbf{F}(t)\bar{\mathbf{Q}}(t) + \bar{\mathbf{Q}}(t)\mathbf{F}^{\mathsf{T}}(t) + \mathbf{Q}_s(t)}, \quad t_{k-1} \leq t \leq t_k$$

  with $\bar{\mathbf{Q}}(t_{k-1}) = \mathbf{0}$ and $\mathbf{Q}_{k-1} = \bar{\mathbf{Q}}(t_k)$.
- Important distinction: $\mathbf{Q}_s$ is a spectral density matrix and $\mathbf{Q}_{k-1}$ is a covariance matrix with different units.
- A spectral density matrix is converted to a covariance matrix through multiplication by the Dirac delta function, $\delta(t-\tau)$ with units of 1/time.

Given

$$\boxed{\mathbf{F}(t), \mathbf{G}(t), \mathbf{Q}_s(t), \mathbf{u}(t), \mathbf{x}_{k-1}}$$

Integrate over $t_{k-1} \le t \le t_k$

$$\boxed{\begin{aligned} \dot{\bar{\mathbf{Q}}}(t) &= \mathbf{F}(t)\bar{\mathbf{Q}}(t) + \bar{\mathbf{Q}}(t)\mathbf{F}^{\mathsf{T}}(t) + \mathbf{Q}_s(t) \\ \dot{\Phi}(t, t_{k-1}) &= \mathbf{F}(t)\Phi(t, t_{k-1}) \end{aligned}} \quad \text{with} \quad \begin{aligned} &\bar{\mathbf{Q}}(t_{k-1}) = \mathbf{0},\ \Phi(t_{k-1}, t_{k-1}) = \mathbf{I} \\ &\mathbf{Q}_{k-1} = \bar{\mathbf{Q}}(t_k),\ \Phi_{k-1} = \Phi(t_k, t_{k-1}) \end{aligned}$$

Map the state forward via

$$\boxed{\mathbf{x}_k = \Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}},$$

where the effect of the input $\mathbf{u}(t)$ over the time interval is computed via

$$\mathbf{u}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\mathbf{G}(\tau)\mathbf{u}(\tau)d\tau$$

and $\mathbf{w}_{k-1}$ is a random input with

$$E\{\mathbf{w}_{k-1}\} = \mathbf{0} \text{ and } E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^{\mathsf{T}}\} = \mathbf{Q}_{k-1}$$

END MODULE