APPLIED KALMAN FILTERING

The State Estimation Error Covariance Matrix and Propagation of State Errors

Dr. Robert H. Bishop, P.E.

- Recall that 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
- The state vector and process noise are vectors whose elements are random variables
- The process noise is assumed to be uncorrelated in time (a white noise process)
- In the discrete time equivalent the process noise is assumed uncorrelated from observation time to observation time (a white noise sequence)

$$\boxed{\mathbf{x}_k = \Phi_{k-1}\mathbf{x}_{k-1} + \mathbf{u}_{k-1} + \mathbf{w}_{k-1}} \tag{2}$$

where $\Phi_{k-1} := \Phi(t_k, t_{k-1})$

- The state estimate at any time t is given by

$$\hat{\mathbf{x}}(t) = E\{\mathbf{x}(t)\}$$

- It follows from the continuous-time mathematical model that

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{u}(t)} \tag{3}$$

  since $E\{\mathbf{w}(t)\} = \mathbf{0}$ and $E\{\mathbf{u}(t)\} = \mathbf{u}(t)$.

- The error in the estimate of a state vector is the difference between the estimated state and the true state

$$\boxed{\mathbf{e}(t) := \mathbf{x}(t) - \hat{\mathbf{x}}(t)} \tag{4}$$

  Note that $E\{\mathbf{e}(t)\} = E\{\mathbf{x}(t)\} - \hat{\mathbf{x}}(t) = \mathbf{0}$.

- Computing $\dot{\mathbf{e}}(t)$ using Eqs. 1 and 3 yields

$$\boxed{\dot{\mathbf{e}}(t) = \mathbf{F}(t)\mathbf{e}(t) + \mathbf{w}(t)} \tag{5}$$

- The covariance $\mathbf{P}(t)$ of $\mathbf{e}(t)$ given by

$$\boxed{\mathbf{P}(t) = E\{\mathbf{e}(t)\mathbf{e}^{\mathsf{T}}(t)\}} \tag{6}$$

provides a statistical measure of the uncertainty in the state estimation error

- The covariance matrix of $\mathbf{e}(t) \in \Re^{n}$ is a symmetric matrix $\mathbf{P}(t) = \mathbf{P}^{\mathsf{T}}(t) \in \Re^{n\times n}$
- The trace of $\mathbf{P}(t)$ is the length of the mean square state estimation error

$$\mathrm{trace}\,(\mathbf{P}(t)) = E\{\mathbf{e}^{\mathsf{T}}(t)\mathbf{e}(t)\} = E\{\|\mathbf{e}(t)\|^{2}\}$$

- The off-diagonal terms of $\mathbf{P}(t)$ are indicators of cross-correlation between the elements of $\mathbf{e}(t)$.

- Recall that the solution to Eq. 1 is given by

$$\mathbf{x}(t) = \Phi(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{u}(\tau)d\tau + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{w}(\tau)d\tau \tag{7}$$

- Then with $\hat{\mathbf{x}}(t) = E\{\mathbf{x}(t)\}$, and $\hat{\mathbf{x}}(t)$ given above, it follows that

$$\hat{\mathbf{x}}(t) = \Phi(t, t_0)\hat{\mathbf{x}}_0 + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{u}(\tau)d\tau \tag{8}$$

  where we employ the assumption that $E\{\mathbf{w}(t)\} = \mathbf{0}, \forall\ t \geq t_0$.
- Using Eqs. 7 and 8 and computing $\mathbf{e}(t)$ yields

$$\boxed{\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t) = \Phi(t, t_0)\mathbf{e}_0 + \int_{t_0}^{t} \Phi(t, \tau)\mathbf{w}(\tau)d\tau} \tag{9}$$

- With $\mathbf{e}(t)$ above, we can derive a relationship for $\mathbf{P}(t)$ in Eq. 6 and take the time-derivative to obtain $\dot{\mathbf{P}}(t)$ describing the rate of change of the state estimation error covariance in time.

- Computing $E\{\mathbf{e}(t)\mathbf{e}^{\mathsf{T}}(t)\}$ yields

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

  where we employ the assumption that $\mathbf{w}(t)$ is a white noise process uncorrelated with $\mathbf{e}(t_0)$.
- With $\mathrm{E}\left\{\mathbf{w}(t)\mathbf{w}^{\mathsf{T}}(\tau)\right\} = \mathbf{Q}_s(t)\delta(t-\tau), \forall t,\tau$, it follows that

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

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

$$\boxed{\mathbf{P}(t) = \Phi(t,t_0)\mathbf{P}_0\Phi^{\mathsf{T}}(t,t_0) + \int_{t_0}^{t}\Phi(t,\tau)\mathbf{Q}_s(\tau)\Phi^{\mathsf{T}}(t,\tau)d\tau} \tag{10}$$

- Recalling that $\dot{\Phi}(t,\tau) = \mathbf{F}(t)\Phi(t,\tau)$, $\mathbf{P}_0$ is constant and using Leibniz's Rule, we compute the time derivative of $\mathbf{P}(t)$ in Eq. 10 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

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

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

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

- Substituting $\mathbf{P}(t)$ in Eq. 10 yields

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

  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 $\mathbf{P}(t) = \mathbf{P}^{\mathsf{T}}(t) > 0, \forall t \geq t_0$

- The state estimate at any time t_k is given by

$$\hat{\mathbf{x}}_k = E\{\mathbf{x}_k\}$$

- It follows from the discrete-time mathematical model that

$$\boxed{\hat{\mathbf{x}}_k = \Phi_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{u}_{k-1}} \tag{11}$$

  since $E\{\mathbf{w}_{k-1}\} = \mathbf{0}$ and $E\{\mathbf{u}_{k-1}\} = \mathbf{u}_{k-1}$.

- The error in the estimate of a state vector is the difference between the estimated state and the true state

$$\boxed{\mathbf{e}_k := \mathbf{x}_k - \hat{\mathbf{x}}_k}$$

  where $\mathbf{x}_k := \mathbf{x}(t_k)$ and $\hat{\mathbf{x}}_k := \hat{\mathbf{x}}(t_k)$. Note that $E\{\mathbf{e}_k\} = E\{\mathbf{x}_k\} - \hat{\mathbf{x}}_k = \mathbf{0}$.

- Computing $\mathbf{e}_k$ using Eqs. 2 and 11 yields

$$\boxed{\mathbf{e}_k = \Phi_{k-1}\mathbf{e}_{k-1} + \mathbf{w}_{k-1}} \tag{12}$$

- The covariance $\mathbf{P}_k$ of $\mathbf{e}_k$ given by

$$\boxed{\mathbf{P}_k = E\{\mathbf{e}_k \mathbf{e}_k^T\}} \tag{13}$$

  provides a statistical measure of the uncertainty in the state estimation error
- The covariance matrix of $\mathbf{e}_k \in \Re^n$ is a symmetric matrix $\mathbf{P}_k = \mathbf{P}_k^T \in \Re^{n \times n}$
- The trace of $\mathbf{P}_k$ is the length of the mean square state estimation error

$$\mathrm{trace}\left(\mathbf{P}_k\right) = E\{\mathbf{e}_k^T \mathbf{e}_k\} = E\{\|\mathbf{e}_k\|^2\}$$

- The off-diagonal terms of $\mathbf{P}_k$ are indicators of cross-correlation between the elements of $\mathbf{e}_k$.

- Computing $\mathbf{P}_k$ in Eq. 13 using Eq. 12 yields

$$\mathbf{P}_k = \Phi_{k-1} E\{\mathbf{e}_{k-1}\mathbf{e}_{k-1}^T\}\Phi_{k-1}^T + \Phi_{k-1} E\{\mathbf{e}_{k-1}\mathbf{w}_{k-1}^T\} + E\{\mathbf{w}_{k-1}\mathbf{e}_{k-1}^T\}\Phi_{k-1}^T + E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T\}$$

- The estimation error $\mathbf{e}_{k-1}$ and the noise $\mathbf{w}_{k-1}$ are uncorrelated (a consequence of the fact that $\mathbf{w}_{k-1}$ is a white sequence), hence $E\{\mathbf{e}_{k-1}\mathbf{w}_{k-1}^T\} = \mathbf{0}$ and $E\{\mathbf{w}_{k-1}\mathbf{e}_{k-1}^T\} = \mathbf{0}$.
- With $\mathbf{P}_{k-1} = E\{\mathbf{e}_{k-1}\mathbf{e}_{k-1}^T\}$ and $\mathbf{Q}_{k-1} = E\{\mathbf{w}_{k-1}\mathbf{w}_{k-1}^T\}$, it follows that

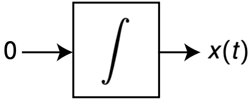
$$\boxed{\mathbf{P}_k = \Phi_{k-1}\mathbf{P}_{k-1}\Phi_{k-1}^T + \mathbf{Q}_{k-1}} \qquad k = 1, 2, \cdots$$

  with the initial condition $\mathbf{P}_0 = \mathbf{P}_0^T > 0$ given and $\mathbf{Q}_{k-1} = \mathbf{Q}_{k-1}^T \geq 0$ given
- Since $\mathbf{P}_0 = \mathbf{P}_0^T > 0$ and $\mathbf{Q}_{k-1} = \mathbf{Q}_{k-1}^T \geq 0$, it follows that $\mathbf{P}_k = \mathbf{P}_k^T > 0$, $\forall k \geq 1$

- A random constant is a non-dynamic quantity with a random constant amplitude modeled via

$$\boxed{\dot{x}(t) = 0}$$

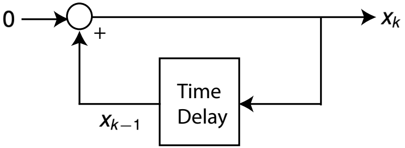
- The output of an integrator with zero input and a random initial condition $x(t_0) = x_0$



- The corresponding discrete process is described by

$$\boxed{x_k = x_{k-1}}$$

  for $k = 1,\ 2,\ \cdots$
- Output of a feedback loop with a time delay with zero input and a random initial condition x_0



- A random walk process is the output of a integrator when the inputs are uncorrelated signals modeled via

$$\boxed{\dot{x}(t) = w(t)}$$

where $E\{w(t)\} = 0$, $E\{w(t)w(\tau)\} = q\delta(t-\tau)$, and q given

- The output of an integrator with uncorrelated inputs and zero initial conditions $x(t_0) = 0$

$$w(t) \longrightarrow \boxed{\int} \longrightarrow x(t)$$

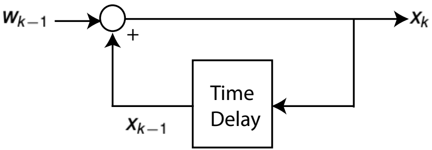
UNIVERSITY of SOUTH FLORIDA

- The corresponding discrete random walk sequence is described by

$$\boxed{x_k = x_{k-1} + w_{k-1}}$$

where $E\{w_{k-1}\} = 0$, $E\{w_{k-1}w_{j-1}\} = Q_{k-1}\delta_{kj}$ where $Q_{k-1} = q\,(t_k - t_{k-1})$

- The output of a feedback loop with a time delay with uncorrelated inputs and zero initial conditions $x_0 = 0$

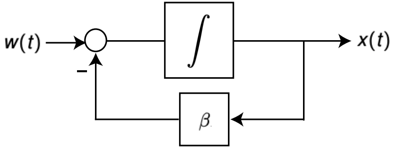


- An exponentially-correlated random variable (ECRV) is the output of a feedback loop when the inputs are uncorrelated signals modeled via

$$\boxed{\dot{x}(t) = -\beta x(t) + w(t)}$$

where $E\{w(t)\} = 0$, $E\{w(t)w(\tau)\} = q\delta(t-\tau)$, $q > 0 \in \Re$ and $\beta > 0 \in \Re$ are given

- The time constant of the individual ECRV is given by $1/\beta$
- The output of a feedback loop with uncorrelated inputs and random initial conditions $x(t_0) = x_0$



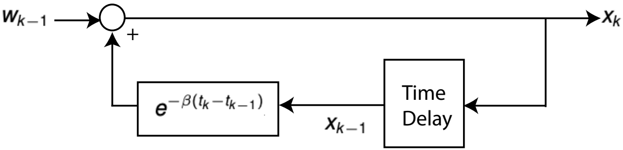
- The corresponding discrete ECRV is described by

$$\boxed{x_k = e^{-\beta(t_k - t_{k-1})} x_{k-1} + w_{k-1}}$$

for $k = 1,\ 2,\ \cdots$ where $E\{w_{k-1}\} = 0$, $E\{w_{k-1} w_{j-1}\} = Q_{k-1}\delta_{kj}$ where

$$Q_{k-1} = \frac{q}{2\beta}\left(1 - e^{-2\beta(t_k - t_{k-1})}\right)$$

- The output of a feedback loop with a time delay with uncorrelated inputs and zero initial conditions $x_0 = 0$



- Consider the system described by a linear, lumped parameter first-order vector-matrix differential equation

$$\boxed{\dot{\mathbf{x}}(t) = \mathbf{F}(t)\mathbf{x}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{w}(t)} \tag{14}$$

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

- Generally we assume that $\mathbf{w}(t)$ is a zero-mean uncorrelated random noise process. What if some components of the process noise are not zero-mean white noise processes?
- We can augment the state vector to account for correlated disturbances
- Suppose $\mathbf{w}(t)$ is composed of some correlated quantities $\mathbf{w}_1(t) \in \Re^r$ where $r \leq n$ and uncorrelated quantities $\mathbf{w}_2(t) \in \Re^n$ such that

$$\mathbf{w}(t) = \mathbf{\Gamma}\mathbf{w}_1(t) + \mathbf{w}_2(t)$$

with $\mathbf{\Gamma} \in \Re^{n \times r}$ and where $\mathbf{w}_1(t)$ can be modeled via

$$\boxed{\dot{\mathbf{w}}_1(t) = \mathbf{F}_1(t)\mathbf{w}_1(t) + \mathbf{w}_3(t)}$$

where $\mathbf{w}_3(t) \in \Re^r$ is a zero-mean, white noise process and $\mathbf{F}_1(t) \in \Re^{r \times r}$ is given

- Let

$$\mathbf{z}(t) = \begin{pmatrix} \mathbf{x}(t) \\ \mathbf{w}_1(t) \end{pmatrix} \in \Re^{n+r}$$

- Then we can write

$$\boxed{\dot{\mathbf{z}}(t) = \bar{\mathbf{F}}(t)\mathbf{z}(t) + \bar{\mathbf{G}}(t)\mathbf{u}(t) + \bar{\mathbf{w}}(t)}$$

where $\mathbf{z}(t) \in \Re^{n+r}$ is the augmented state vector, $\mathbf{u}(t) \in \Re^{p}$ is a vector of known external inputs, $\bar{\mathbf{w}}(t)$ is the augmented process noise with

$$\bar{\mathbf{F}}(t) = \begin{bmatrix} \mathbf{F}(t) & \mathbf{\Gamma} \\ \mathbf{0} & \mathbf{F}_1(t) \end{bmatrix} \in \Re^{(n+r)\times(n+r)}, \ \bar{\mathbf{G}}(t) = \begin{bmatrix} \mathbf{G}(t) \\ \mathbf{0} \end{bmatrix} \in \Re^{(n+r)\times p}, \ \bar{\mathbf{w}}(t) = \begin{pmatrix} \mathbf{w}_2(t) \\ \mathbf{w}_3(t) \end{pmatrix} \in \Re^{n+r}$$

- The augmented process noise $\bar{\mathbf{w}}(t)$ is a zero-mean white noise process, as desired

END MODULE