APPLIED KALMAN FILTERING

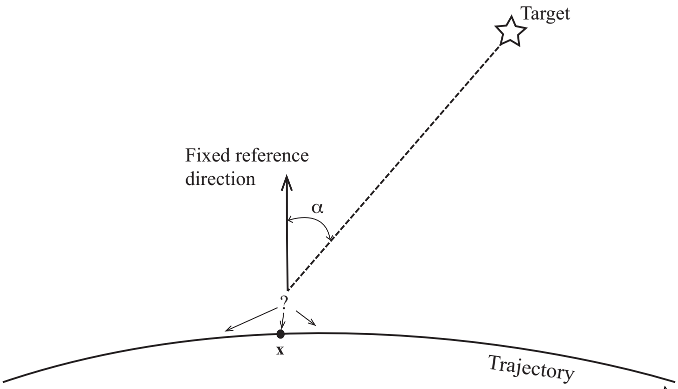
Kalman Filter

Dr. Robert H. Bishop, P.E.

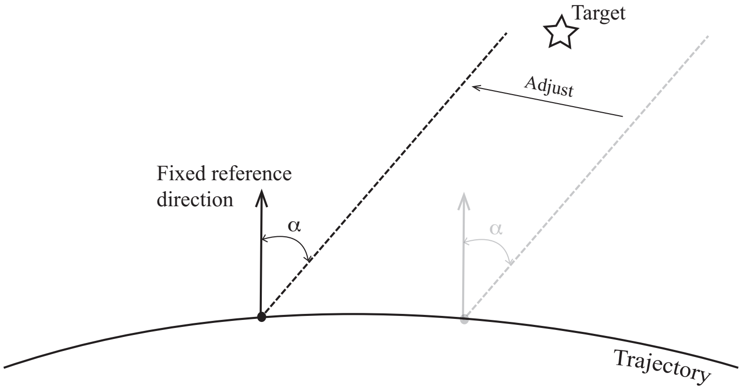
UNIVERSITY of SOUTH FLORIDA

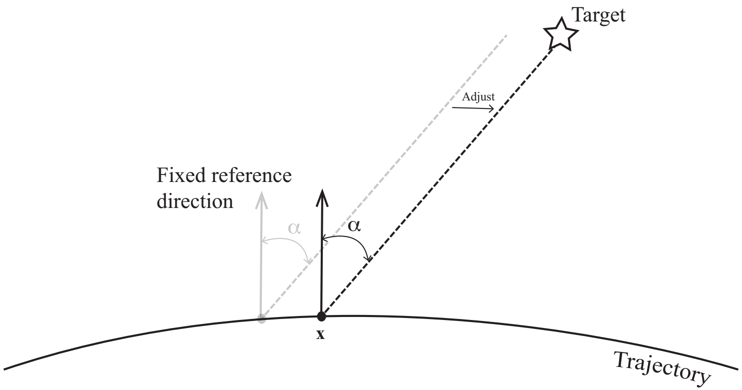
- Objective: Determine the best estimate of the spacecraft position along a known trajectory based on one noise-free measurement
- Assumptions:
  - Trajectory is know perfectly
  - Position on the trajectory is unknown
  - No initial estimate of position, denoted by $\mathbf{x}$
  - Only one position $\mathbf{x}$ exists
- Estimate the position in geometric alignment with the measurement
- This has nothing to do with statistics

Measure the angle between a fixed reference and a target



Adjust the position estimate to try to align with the measurement

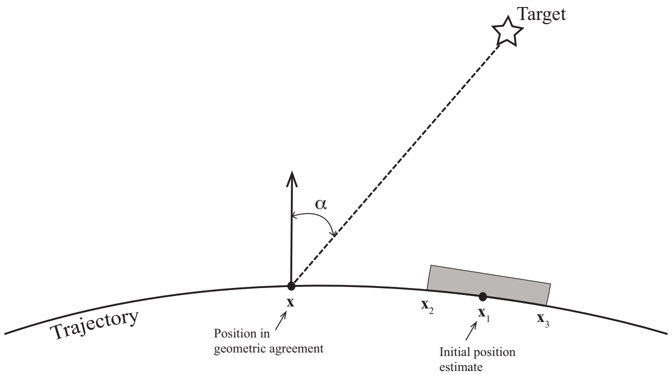




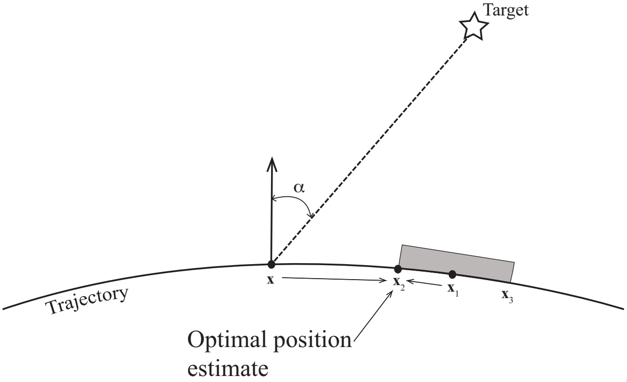
- Since the measurement is error-free, then **x** is the true position

- Objective: Determine the best estimate of the spacecraft position along a known trajectory based on one noisy measurement and an initial position estimate with uniformly distributed errors
- Assumptions:
  - Trajectory is known perfectly
  - Exact position on the trajectory is unknown
  - An initial position estimate exists, denoted by $\mathbf{x}_1$, with a uniformly distributed error
  - Only one actual position $\mathbf{x}$ exists and is not between $\mathbf{x}_2$ and $\mathbf{x}_3$
- $\mathbf{x}_2$ is the best estimate since it agrees more closely with the measurement than any other point in the shaded region.
- Initial estimate is moved from $\mathbf{x}_1$ towards $\mathbf{x}$, but not all the way

Measure the angle between a fixed reference and a target



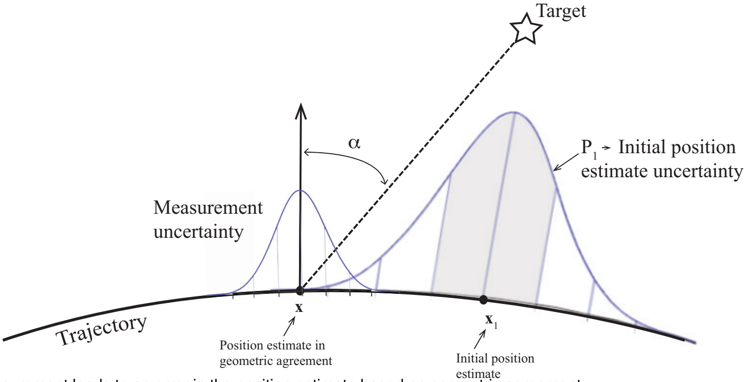
The position estimate is updated to split the difference between the initial estimate and the position in geometric agreement with the measurement



UNIVERSITY of SOUTH FLORIDA

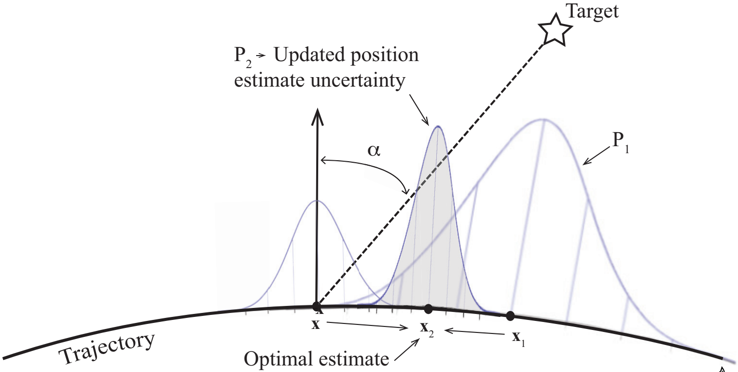
- Objective: Determine the best estimate of the spacecraft position along a known trajectory based on one noisy measurement and an initial position estimate with normally distributed errors
- Assumptions:
  - Trajectory is known perfectly
  - Exact position on the trajectory is unknown
  - An initial position estimate exists, denoted by $\mathbf{x}_1$, with a normally distributed error
- $\mathbf{x}_2$ is the optimal estimate based on the principle of statistical weighting.
- The further $\mathbf{x}$ is from $\mathbf{x}_1$, the more likely it is that the actual position lies between $\mathbf{x}$ and $\mathbf{x}_1$.
- The higher the confidence in the measurement, the closer $\mathbf{x}_2$ is to $\mathbf{x}$.
- The statistical model of the estimation error in the estimate $\mathbf{x}_2$ changes.

Measure the angle between a fixed reference and a target



Note: The noisy measurement leads to an error in the position estimate based on geometric agreement.

Both the position estimate and the position estimate uncertainty are updated according to the principle of statistical weighting



- From a statistical point of view, the situation has been improved by this one measurement
- However, it is possible that the actual estimation error has increased by this single measurement
- To build confidence, it is necessary to obtain a sequence of measurements
- Each incorporation of the measurement will, on the average, decrease the estimation error

- Estimate the distance to the wall, x, using a noisy laser ranger
- Model the laser ranger output at time t_i as

$$\boxed{y_i = x + v_i}$$

where v_i is a white noise sequence with $E\{v_i\} = 0$, $E\{v_i^2\} = R$, $\forall i$, and $E\{v_i v_j\} = 0$ for $i \neq j$

- Suppose we take n measurements at $t_1, t_2, \cdots, t_n$, denoted by $y_1, y_2, \cdots, y_n$, respectively, Denote the estimate of x at t_n as $\hat{x}_n$. Estimate the distance with a simple average as

$$\hat{x}_n = \frac{1}{n}\sum_{i=1}^{n} y_i$$

- Note that with the mathematical model $\boxed{y_i = x + v_i}$ it follows that

$$\hat{x}_n = \frac{1}{n}\sum_{i=1}^{n} y_i = x + \frac{1}{n}\sum_{i=1}^{n} v_i = x + \nu_n$$

  where we define $\nu_n := \frac{1}{n}\sum_{i=1}^{n} v_i$. We have $E\{\nu_n\} = 0$ and $E\{\nu_n^2\} = R/n$

- The averaging algorithm has reduced the noise strength by a factor of $1/n \rightarrow$ more measurements reduces the estimation error, in other words, the estimate is closer to the truth

- Suppose we take an additional measurement y_{n+1} at t_{n+1}. If we have saved (non-recursive) all the previous n measurements, we can compute the estimate via

$$\hat{x}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} y_i$$

- Let's re-think this:

$$\hat{x}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n} y_i + \frac{1}{n+1} y_{n+1} = \frac{n}{n+1} \hat{x}_n + \frac{1}{n+1} y_{n+1}$$

- If we define $K_1 := \frac{n}{n+1}$ and $K_2 := \frac{1}{n+1}$ we have

$$\boxed{\hat{x}_{n+1} = K_1 \hat{x}_n + K_2 y_{n+1}}$$

- This is an example of Principle of Statistical Weighting–except in this simple example K_1 and K_2 are fixed. In general, if we have more confidence in the estimate $\hat{x}_n$ we select K_1 larger than K_2, and vice versa

- We can rewrite $\hat{x}_{n+1} = K_1 \hat{x}_n + K_2 y_{n+1}$ as

$$\boxed{\hat{x}_{n+1} = \hat{x}_n + K_n \left(y_{n+1} - \hat{y}_{n+1} \right)} \quad n = 1,\ 2,\ \cdots$$

where $K_n = 1/(n+1)$ and $\hat{y}_{n+1}$ is our estimate of the measurement which is derived from a model of the sensor. Here we let $\hat{y}_{n+1} = \hat{x}_n$, the last available state estimate and $\hat{x}_1$ is given

<u>Recursive</u> – no need to save any past measurements
<u>Real-time</u> – updated estimate is computed with each measurement

- Define the estimation errors and estimation error covariances at t_n and t_{n+1} as

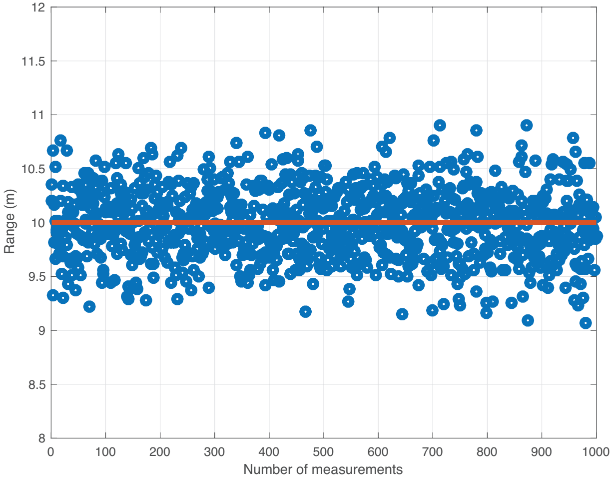
$$e_n = x - \hat{x}_n,\ e_{n+1} = x - \hat{x}_{n+1} \quad \text{and} \quad P_n = E\{e_n^2\},\ P_{n+1} = E\{e_{n+1}^2\}$$

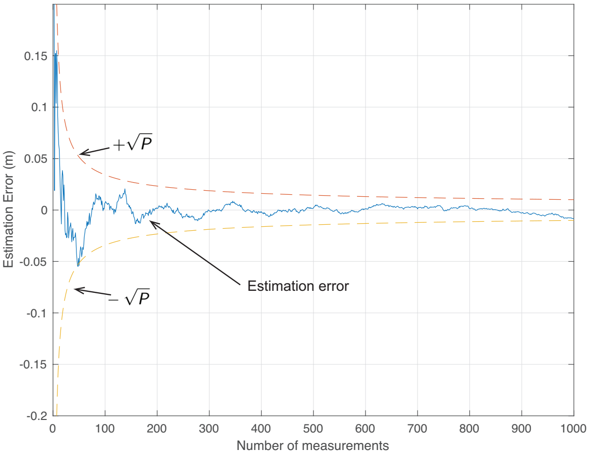
- For $n = 1,\ 2,\ \cdots$, it follows that

$$\boxed{P_{n+1} = \left(\frac{n}{n+1} \right)^2 P_n + \left(\frac{1}{n+1} \right)^2 R} \quad \text{with } P_1 \text{ given}$$

- In general, we would use a sensor mathematical model for $\hat{y}_{n+1}$ and K_n would be computed via Kalman filtering theory

Simulated data–the measurements with $R = 0.1$ m^2



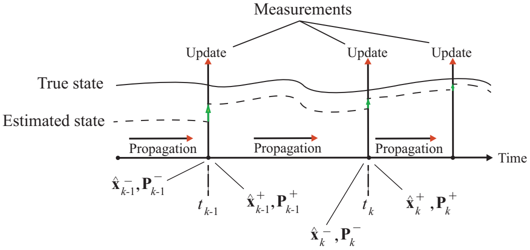

$+\sqrt{P}$
$-\sqrt{P}$
Estimation error
Estimation Error (m)
Number of measurements
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
0
100
200
300
400
500
600
700
800
900
1000

- **What?** A design and analysis of a recursive real-time algorithm to provide state estimates
- **Why?** Estimation errors drive performance.

> Provide real-time estimates of the state of the system to other subsystems at required rates in required frames in the presence of significant modeling uncertainties and sensor errors

- **Let's ease into it** $\cdots$ Use the Principle of Statistical Weighting to develop the Kalman filter (KF) architecture for linear systems

- The Kalman filter has two main stages: propagation and update



- The update stage using sensor measurements from multiple sensors can occur asynchronously and at varying time intervals
- At each measurement update, the state estimate and the state estimation error covariance will exhibit a discrete update
- On average, we expect the state estimate to converge to the true state and the state estimation error covariance to reflect the true uncertainty in the state estimate

- Begin with the Principle of Statistical Weighting

$$\boxed{\hat{\mathbf{x}}_k^+ = \mathbf{K}_{1,k}\hat{\mathbf{x}}_k^- + \mathbf{K}_{2,k}\mathbf{y}_k}$$

- We have the current state estimate and latest external measurement
- How do we choose $\mathbf{K}_{1,k}$ and $\mathbf{K}_{2,k}$?
- Theory is based on linear mathematical models for both dynamics and sensors $\leftarrow$ typically not the case in the real-world
- Define the estimation errors:

$$\boxed{\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-} \text{ (\textit{apriori}) and } \boxed{\mathbf{e}_k^+ = \mathbf{x}_k - \hat{\mathbf{x}}_k^+} \text{ (\textit{aposteriori})}$$

- We need to compute $\mathbf{K}_{1,k}$ and $\mathbf{K}_{2,k}$ so that the KF 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

- After a little algebra ....

$$\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^- \qquad \mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k$$

$$\mathbf{e}_k^+ = \mathbf{x}_k - \hat{\mathbf{x}}_k^+ \qquad \boxed{\hat{\mathbf{x}}_k^+ = \mathbf{K}_{1,k}\hat{\mathbf{x}}_k^- + \mathbf{K}_{2,k}\mathbf{y}_k}$$

$$\mathbf{e}_k^+ = \mathbf{K}_{1,k}\mathbf{e}_k^- + (\mathbf{I} - \mathbf{K}_{1,k} - \mathbf{K}_{2,k}\mathbf{H}_k)\,\mathbf{x}_k - \mathbf{K}_{2,k}\mathbf{v}_k$$

- Taking the expectation yields

$$E\{\mathbf{e}_k^+\} = \mathbf{K}_{1,k} E\{\mathbf{e}_k^-\} + (\mathbf{I} - \mathbf{K}_{1,k} - \mathbf{K}_{2,k}\mathbf{H}_k)\, E\{\mathbf{x}_k\} - \mathbf{K}_{2,k} E\{\mathbf{v}_k\}$$

- We know $E\{\mathbf{v}_k\} = \mathbf{0}$. Suppose $E\{\mathbf{e}_k^-\} = \mathbf{0}$ (more on this later)
- Since we want $E\{\mathbf{e}_k^+\} = \mathbf{0}$ for all $\mathbf{x}_k$, we require $\mathbf{K}_{1,k} = \mathbf{I} - \mathbf{K}_{2,k}\mathbf{H}_k$
- The result is that

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

where $\mathbf{K}_k := \mathbf{K}_{2,k}$

UNIVERSITY of SOUTH FLORIDA

- After a little more algebra ....

$$\mathbf{e}_k^+ = \cancelto{[\mathbf{I}-\mathbf{K}_k\mathbf{H}_k]}{\mathbf{K}_{1,k}}\mathbf{e}_k^- + (\mathbf{I} - \cancelto{\mathbf{0}}{\mathbf{K}_{1,k}} - \mathbf{K}_{2,k}\mathbf{H}_k)\,\mathbf{x}_k - \cancelto{\mathbf{K}_2}{\mathbf{K}_{2,k}}\mathbf{v}_k$$

$$\Downarrow$$

$$\mathbf{e}_k^+ = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\mathbf{e}_k^- + \mathbf{K}_k\mathbf{v}_k$$

- Define the *apriori* and *aposteriori* state estimation covariance matrices:

$$\mathbf{P}_k^- := E\{\mathbf{e}_k^-\mathbf{e}_k^{-\,\mathrm{T}}\} \quad \text{and} \quad \mathbf{P}_k^+ := E\{\mathbf{e}_k^+\mathbf{e}_k^{+\,T}\}$$

- With the estimation error and estimation error covariance defined as above, and with the fact that $E\{\mathbf{e}_k^-\mathbf{v}_k^T\} = \mathbf{0}$ we have

$$\boxed{\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\mathbf{P}_k^-\,[\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]^{\mathrm{T}} + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^{\mathrm{T}}} \leftarrow \textit{Joseph Formula}$$

- Goal is to minimize the expected mean squared error, $\mathcal{J}_k = E\{\mathbf{e}_k^{+\,T}\mathbf{e}_k^+\}$

$$\mathbf{P}_k^+ := E\{\mathbf{e}_k^+\mathbf{e}_k^{+\,T}\} \rightarrow \mathcal{J}_k = tr\left\{E\{\mathbf{e}_k^+\mathbf{e}_k^{+\,T}\}\right\} = tr\;\mathbf{P}_k^+$$

UNIVERSITY of SOUTH FLORIDA

- Search for $\mathbf{K}_k$ to minimize $\mathcal{J}_k = tr\ \mathbf{P}_k^+$, or

$$tr\ \mathbf{P}_k^+ = tr\ \mathbf{P}_k^- - 2tr\ \mathbf{K}_k\mathbf{H}_k\mathbf{P}_k^- + tr\ \mathbf{K}_k \left[\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}\right] \mathbf{K}_k^\mathsf{T}$$

- Approach: Take the partial derivative of $tr\ \mathbf{P}_k^+$ with respect to $\mathbf{K}_k$, set the result to zero and solve for the optimal $\mathbf{K}_k$
- Recall a few really useful relationships

$$\frac{\partial}{\partial \mathbf{A}} tr\mathbf{AB} = \mathbf{B}^\mathsf{T} \quad \text{and} \quad \frac{\partial}{\partial \mathbf{A}} tr\mathbf{ABA}^\mathsf{T} = 2\mathbf{AB}^\mathsf{T}$$

- So we have

$$\frac{\partial}{\partial \mathbf{K}_k} tr\mathcal{J}_k = -2\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + 2\mathbf{K}_k \left[\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k\right] = \mathbf{0}$$

- Optimal gain matrix $\mathbf{K}_k$ is

$$\boxed{\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} \left[\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k\right]^{-1}} \leftarrow \textit{Kalman Gain}$$

- Substituting the Kalman gain $\mathbf{K}_k$ into the Joseph formula yields

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

- At any time, $t_{k-1} \leq t \leq t_k$, the state estimate is given by

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

- It follows from the continuous-time mathematical model that

$$\boxed{\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{u}(t)}, \quad t_{k-1} \leq t \leq t_k \tag{1}$$

  with the initial condition

$$\hat{\mathbf{x}}(t_{k-1}) = \hat{\mathbf{x}}_{k-1}^{+}$$

- It follows from the discrete-time mathematical model that

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

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

- The two ways of propagating the state estimate in Eqs. (1) and (2) are mathematically equivalent

- As before, we define

$$\mathbf{e}_k^- := \mathbf{x}_k - \hat{\mathbf{x}}_k^- \ , \ \mathbf{e}_{k-1}^+ := \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1}^+ \ , \ \mathbf{P}_k^- := E\{\mathbf{e}_k^- \mathbf{e}_k^{-\,\mathsf{T}}\} \ , \ \mathbf{P}_{k-1}^+ := E\{\mathbf{e}_{k-1}^+ \mathbf{e}_{k-1}^{+\,T}\}$$

- The state estimation error is then given by

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

- Note that

$$E\{\mathbf{w}_{k-1}\} = \mathbf{0}$$

  so, if $E\{\mathbf{e}_{k-1}^+\} = \mathbf{0}$ (unbiased update), then $E\{\mathbf{e}_k^-\} = \mathbf{0}$, as desired ✓
- Computing $\mathbf{P}_k^- = E\{\mathbf{e}_k^- \mathbf{e}_k^{-\,\mathsf{T}}\}$ yields

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

  where

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

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

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

UNIVERSITY of SOUTH FLORIDA

Initial conditions	$\hat{\mathbf{x}}_0 = \mathrm{E}\left\{\mathbf{x}(t_0)\right\}$ $\mathbf{P}_0 = \mathrm{E}\left\{(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^{\mathrm{T}}\right\}$
Integrate $t_{k-1} \leq t \leq t_k$	$\dot{\Phi}(t, t_{k-1}) = \mathbf{F}(t)\Phi(t, t_{k-1}),\ \Phi(t_{k-1}, t_{k-1}) = \mathbf{I}$ $\dot{\bar{\mathbf{Q}}}(t) = \mathbf{F}(t)\bar{\mathbf{Q}}(t) + \bar{\mathbf{Q}}(t)\mathbf{F}^{\mathrm{T}}(t) + \mathbf{Q}_s(t),\ \bar{\mathbf{Q}}(t_{k-1}) = \mathbf{0}$ $\Phi_{k-1} := \Phi(t_k, t_{k-1})\ ,\ \mathbf{Q}_{k-1} := \bar{\mathbf{Q}}(t_k)$
Known Input	$\mathbf{u}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau)\mathbf{G}(\tau)\mathbf{u}(\tau)d\tau$
Propagate	$\hat{\mathbf{x}}_k^- = \Phi_{k-1}\hat{\mathbf{x}}_{k-1}^+ + \mathbf{u}_{k-1}$ $\mathbf{P}_k^- = \Phi_{k-1}\mathbf{P}_{k-1}^+\Phi_{k-1}^{\mathrm{T}} + \mathbf{Q}_{k-1}$
Update	$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}}\left[\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k\right]^{-1}$ $\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^-\right)$ $\mathbf{P}_k^+ = \left[\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right]\mathbf{P}_k^-$

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

END MODULE

UNIVERSITY of SOUTH FLORIDA