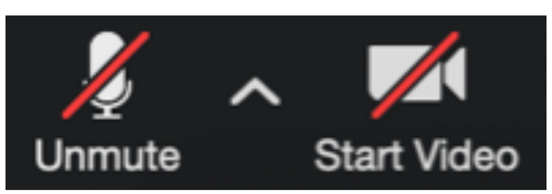
AEROSP 584 - Navigation and Guidance: From Perception to Control



Lectures start at 10:30am EST

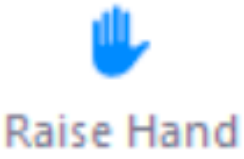
Vasileios Tzoumas

Lecture 10

Slides by Ankit Goel

COLLEGE OF ENGINEERING
AEROSPACE ENGINEERING
UNIVERSITY OF MICHIGAN

To ask questions:



or

Outline

- What is filtering?
  - State estimator

- A little Probability Theory
  - Random vectors
  - Expected value
  - Covariance

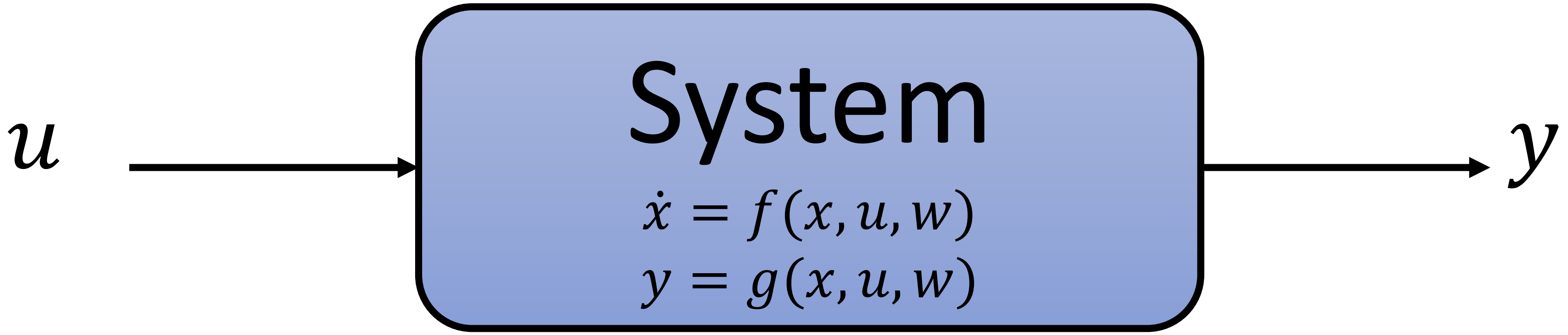
- Optimal State Estimation
  - Optimal Predictor
  - Kalman Filter

Why?

- Navigation: Where are you?
- We are interested in where the vehicle is and where it is going
  - Position + Velocity
- We can measure position and velocity
  - Radar, GPS
- We know the equations of motion, we can integrate them
- We can combine the measurements with the equations of motion

State Estimation

What is State Estimation?



- State estimation is the process of estimating the internal state x of the system using
    - Input u
    - Output y
    - Functions f and g

State Estimation - Example

$$\ddot{x} = u + w$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w \qquad (\star)$$

- To get $x(t)$, integrate $(\star)$, but need to know $x(0)$ and w

- Let the velocity be measured. Then

$$y = \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ \dot{x} \end{bmatrix} + v$$

  - Use y to compensate for the lack of knowledge of $x(0)$ and the noise

Why Use a State Estimator?

UNIVERSITY OF MICHIGAN

- Estimates states that are not directly measured
  - We only measure y
  - We want x

- Reduce the variance of the measured state
  - Suppose you are only interested in the state that you measured
  - Filtering may improve the accuracy of the measured state

- Merge asynchronous measurements
  - A filter can blend various measurements

State Estimation

Unknown
Measured
Computed

- Consider the system

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$
$$y_k = Cx_k + D_2 v_k$$

$x_k \in \mathbb{R}^{l_x}$ is the state
$u_k \in \mathbb{R}^{l_u}$ is the input
$y_k \in \mathbb{R}^{l_y}$ is the measured output
$w_k \in \mathbb{R}^{l_w}$ is the process noise
$v_k \in \mathbb{R}^{l_w}$ is the measurement noise

- The goal of the state estimator is to construct an estimate $\hat{x}_k$ of the state x_k that is best in some sense

Naïve State Estimation – Error Dynamics

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k$$

- Define $e_k \triangleq \hat{x}_k - x_k$. Then,

$$e_{k+1} = Ae_k - D_1 w_k$$

$$e_k = A^k e_0 - \sum_{i=0}^{k} A^{k-i} D_1 w_i$$

- If A is Unstable or Marginally stable, then this method doesn't work
- If A is asymptotically stable, then this method may work
  - Decay of e_0 governed by the eigenvalues of A

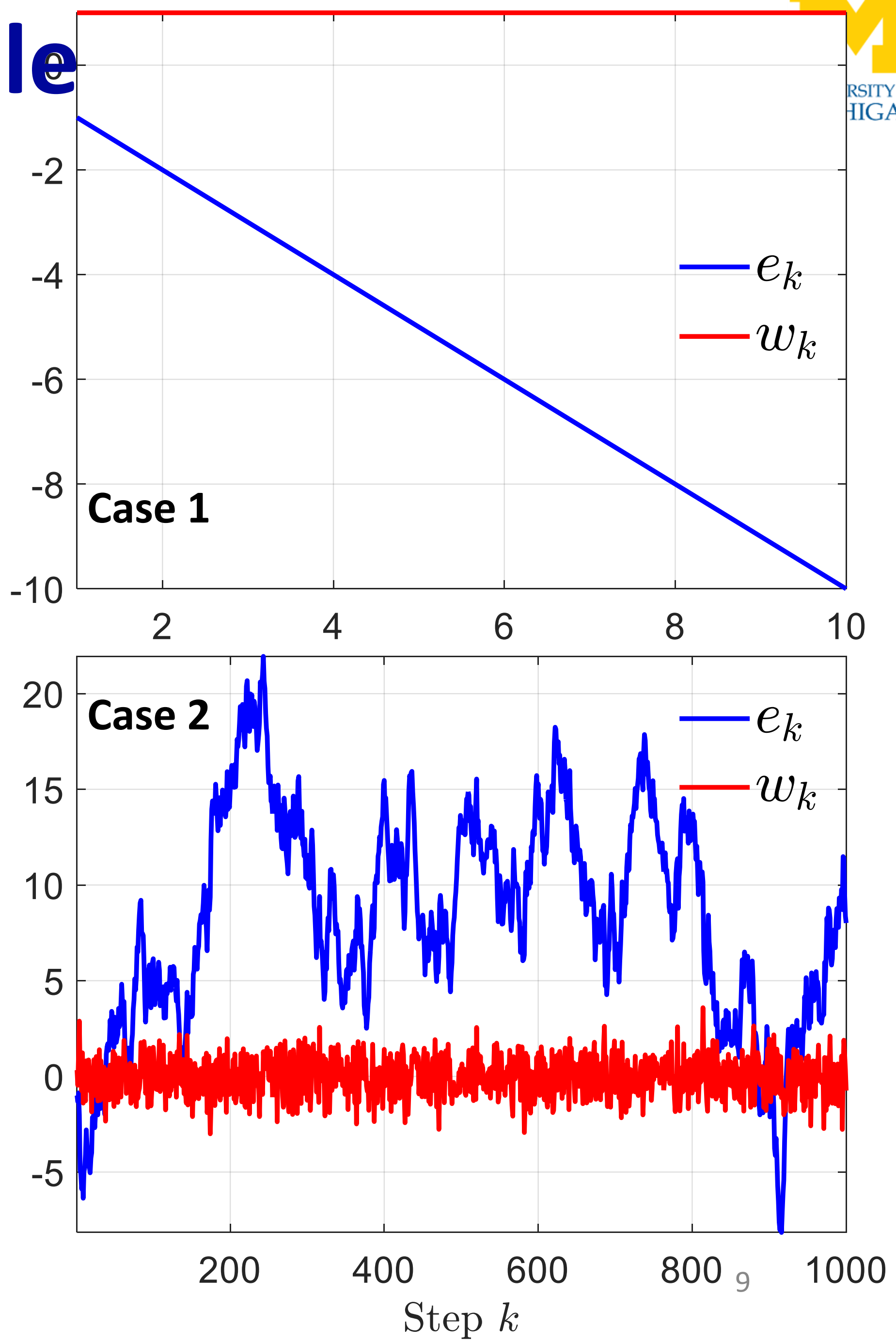
Naïve State Estimation – Example

$$x_{k+1} = x_k + u_k + w_k$$
$$\hat{x}_{k+1} = \hat{x}_k + u_k$$

$$e_k = e_0 - \sum_{i=0}^{k} w_i$$

$x_0 = 1, \hat{x}_0 = 0$

- Case 1 - $w_k = 1$
- Case 2 - $w_k \sim \mathcal{N}(0,1)$



State Estimator

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$
$$y_k = Cx_k + D_2 v_k$$

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K(y_k - C\hat{x}_k)$$

- Define $e_k \triangleq \hat{x}_k - x_k$. Then,

$$e_{k+1} = (A - KC)e_k + KD_2 v_k - D_1 w_k$$

- If (A, C) is observable, then the eigenvalues of $A - KC$ can be placed arbitrarily.

Simple Example

$$x_{k+1} = x_k + u_k + w_k$$
$$y_k = x_k + v_k$$

$$\hat{x}_{k+1} = \hat{x}_k + u_k + K(y_k - \hat{y}_k)$$
$$\hat{y}_k = \hat{x}_k$$

$$e_{k+1} = (1 - K)e_k + Kv_k - w_k$$

Simple Example

How should we choose K?

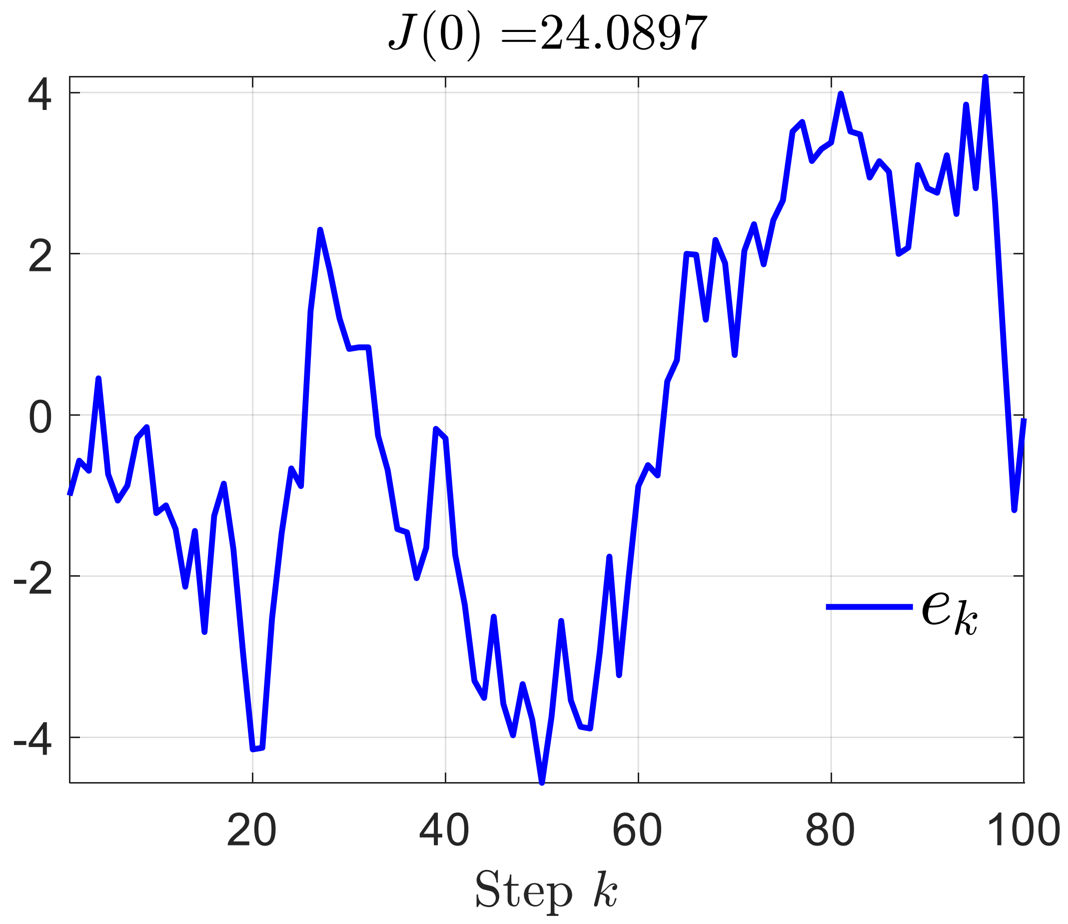
$$e_{k+1} = (1 - K)e_k + Kv_k - w_k$$

Run for 100 steps with

$$x_0 = 1, \hat{x}_0 = 0 \Rightarrow e_0 = 1$$

$$v_k, w_k \sim \mathcal{N}(0,1)$$

$$J(K) = \sum_{i=0}^{100} e_i^2$$



$$e_{k+1} = e_k - w_k$$

Simple Example

How should we choose K?

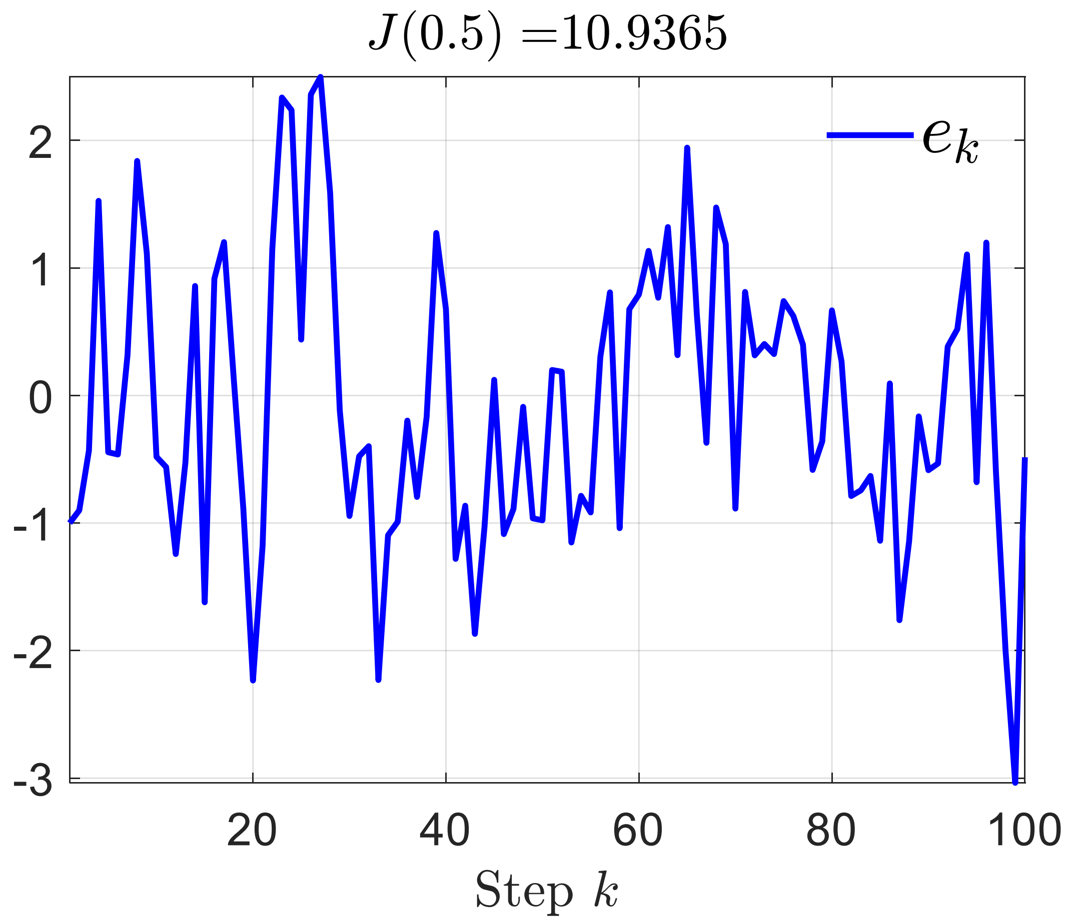
$$e_{k+1} = (1-K)e_k + Kv_k - w_k$$

Run for 100 steps with

$$x_0 = 1, \hat{x}_0 = 0 \Rightarrow e_0 = 1$$

$$v_k, w_k \sim \mathcal{N}(0,1)$$

$$J(K) = \sum_{i=0}^{100} e_i^2$$



$$e_{k+1} = 0.5e_k + 0.5v_k - w_k$$

Simple Example

How should we choose K?

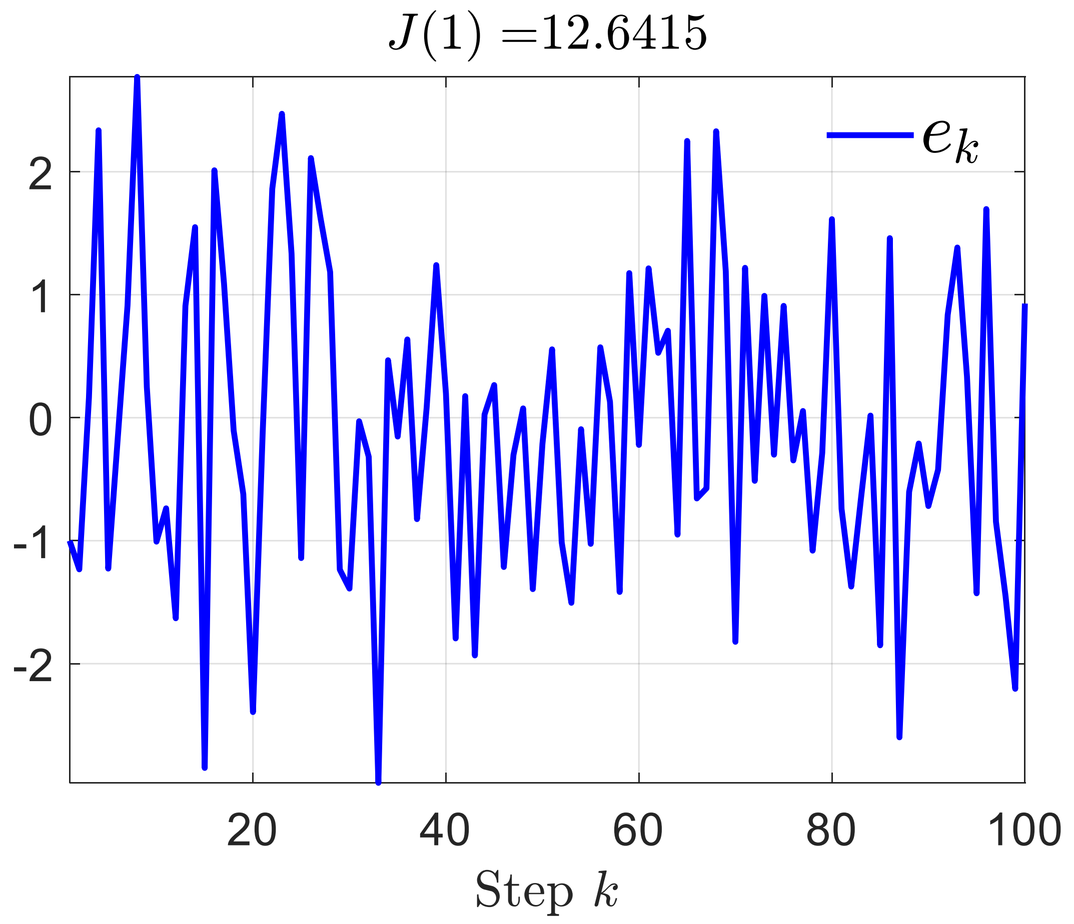
$$e_{k+1} = (1-K)e_k + Kv_k - w_k$$

Run for 100 steps with

$$x_0 = 1, \hat{x}_0 = 0 \Rightarrow e_0 = 1$$

$$v_k, w_k \sim \mathcal{N}(0,1)$$

$$J(K) = \sum_{i=0}^{100} e_i^2$$



$$e_{k+1} = v_k - w_k$$

Simple Example

How should we choose K?

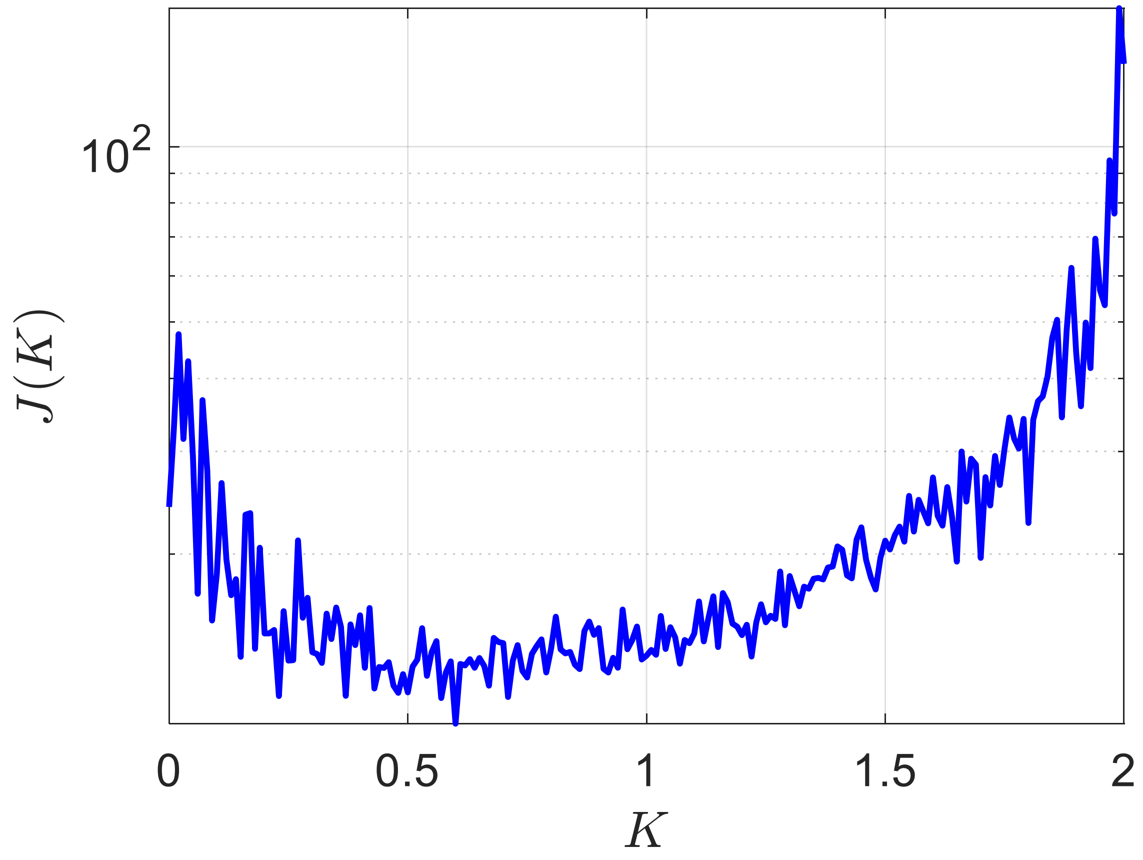
$$e_{k+1} = (1-K)e_k + Kv_k - w_k$$

Run for 100 steps with

$$x_0 = 1, \hat{x}_0 = 0 \Rightarrow e_0 = 1$$

$$v_k, w_k \sim \mathcal{N}(0,1)$$

$$J(K) = \sum_{i=0}^{100} e_i^2$$



$$e_{k+1} = (1-K)e_k + Kv_k - w_k$$

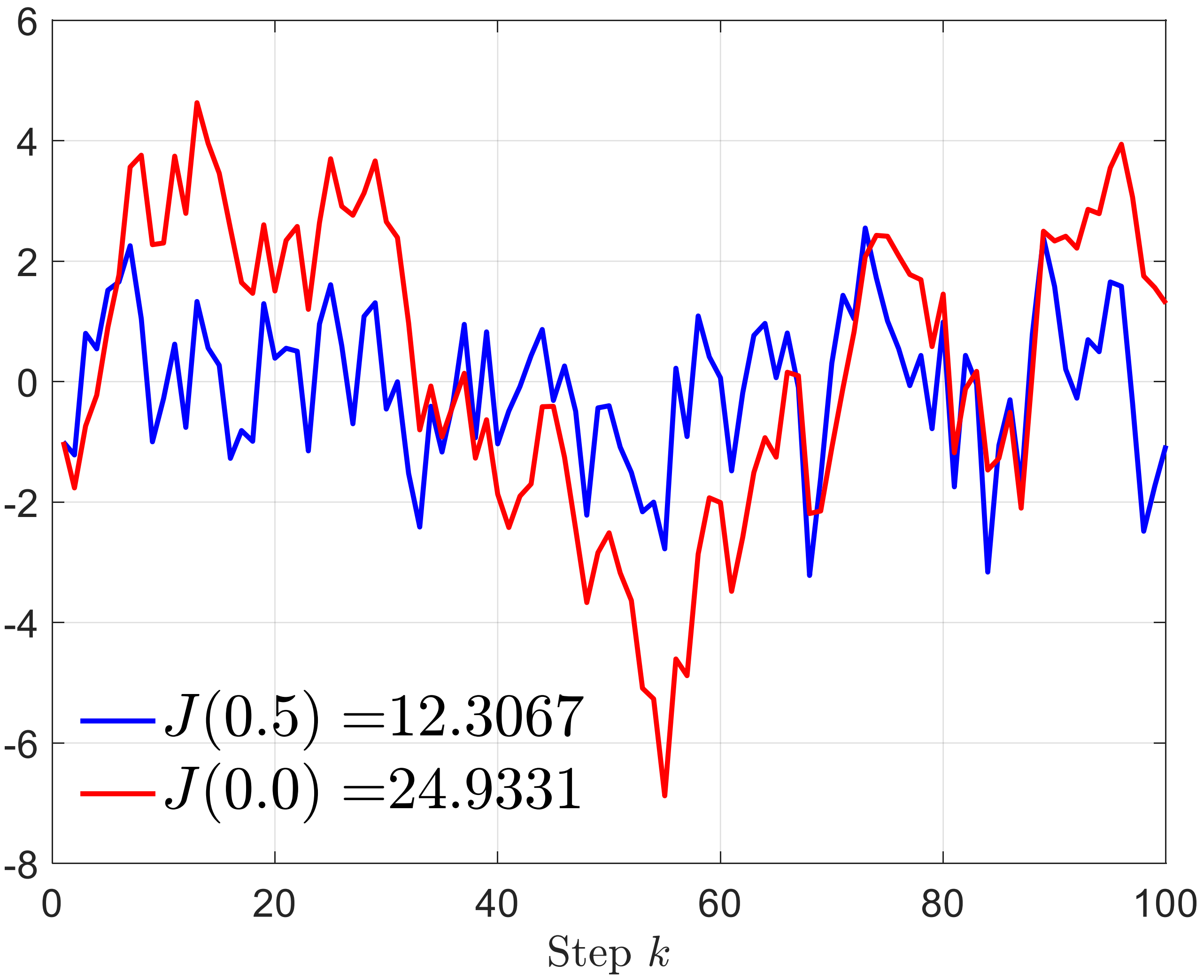
Simple Example

$$x_{k+1} = x_k + u_k + w_k$$
$$y_k = x_k + v_k$$

$$\hat{x}_{k+1} = \hat{x}_k + u_k + K(y_k - \hat{x}_k)$$

$$e_{k+1} = (1-K)e_k + Kv_k - w_k$$

$$J(K) = \sum_{i=0}^{100} e_i^2$$



So far…

- Injecting the measurement can improve the estimate accuracy
- How much to inject?

- We use probability theory to optimize the gain K

Random Vector

- A random vector is a function
$$X: \Omega \to \mathbb{R}^n$$
  - Ω is the sample space
  - $X(\omega) = x \in \mathbb{R}^n$ is the value of the RV associated with the event ω

- The probability density function of a random vector X is the non-negative function $f_X: \mathbb{R}^n \to \mathbb{R}$ that satisfies

$$\Pr(X(\omega) \in \mathcal{D} \subset \mathbb{R}) = \int_{\mathcal{D}} f_X(x) \mathrm{d}x$$

Random Vector

- Expected value of X

$$\mathbb{E}[X] \triangleq \int_{\mathbb{R}^n} x f_X(x) \mathrm{d}x$$

$$\mathbb{E}[AX+b] = \int_{\mathbb{R}^n} (Ax+b) f_X(x) \mathrm{d}x = A\mathbb{E}[X] + b$$

- Covariance of X

$$\begin{aligned}\mathrm{Cov}[X] &\triangleq \mathbb{E}[(X-\mathbb{E}[X])(X-\mathbb{E}[X])^{\mathrm{T}}] \\ &= \int_{\mathbb{R}^n} (X-\mathbb{E}[X])(X-\mathbb{E}[X])^{\mathrm{T}} f_X(x) \mathrm{d}x\end{aligned}$$

$$\begin{aligned}\mathrm{Cov}[AX+b] &= \mathbb{E}[(AX+b-\mathbb{E}[AX+b])(AX+b-\mathbb{E}[AX+b])^{\mathrm{T}}] \\ &= \mathbb{E}[(AX-A\mathbb{E}[X])(AX-A\mathbb{E}[X])^{\mathrm{T}}] \\ &= A\mathbb{E}[(X-\mathbb{E}[X])(X-\mathbb{E}[X])^{\mathrm{T}}]A^{\mathrm{T}} \\ &= A\mathrm{Cov}[X]A^{\mathrm{T}}\end{aligned}$$

Random Vector – Cross-covariance

- Cross-covariance of X and Y

$$
\begin{aligned}
\mathrm{Cov}[X, Y] &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^{\mathrm{T}}] \\
&= \mathbb{E}\left[\left(XY^{\mathrm{T}} + \mathbb{E}[X]\mathbb{E}[Y]^{\mathrm{T}} - \mathbb{E}[X]Y^{\mathrm{T}} - X\mathbb{E}[Y]^{\mathrm{T}}\right)\right] \\
&= \mathbb{E}[XY^{\mathrm{T}}] - \mathbb{E}[X]\mathbb{E}[Y]^{\mathrm{T}}
\end{aligned}
$$

$$
\mathrm{Cov}[X, Y] = \int_{\mathbb{R}^n \times \mathbb{R}^n} (X - \mathbb{E}[X])(Y - \mathbb{E}[Y])^{\mathrm{T}} f_{X,Y}(x, y) \mathrm{d}x \, \mathrm{d}y
$$

$$
\mathrm{Cov}[X] = \mathrm{Cov}[X, X]
$$

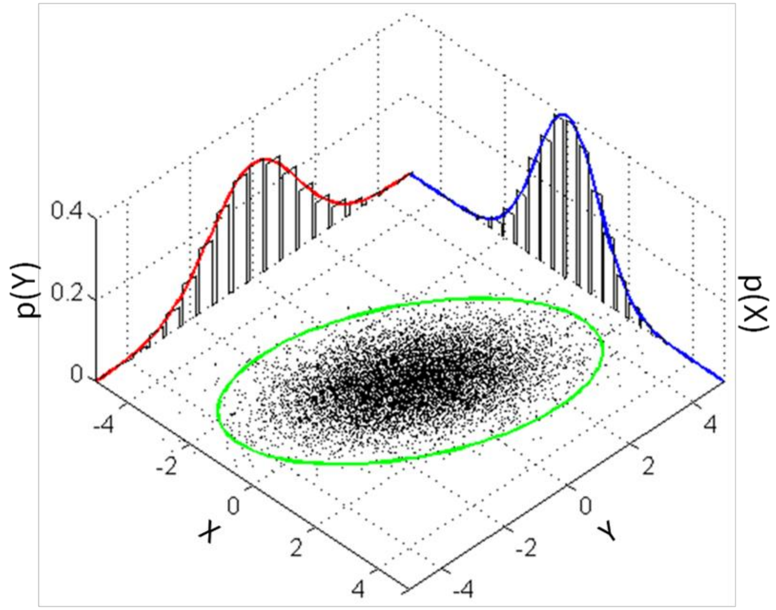
Gaussian Random Vector

- A Gaussian random vector is random vector with pdf

$$f_X(x) = \frac{1}{\sqrt{2\pi \det \Sigma}} \exp \frac{-(x-\mu)^{\mathrm{T}} \Sigma^{-1} (x-\mu)}{2}$$

- Expected value of $X = \mu$
- Covariance of $X = \Sigma > 0$

$$w \sim \mathcal{N}(\mu, \Sigma)$$



Next …

- Optimize the gain K in

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K(y_k - C\hat{x}_k)$$

using the dynamics and the noise properties

A Convenient Reformulation

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + D_1 w_k \\ y_k &= Cx_k + D_2 v_k \end{aligned}$$

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \begin{bmatrix} D_1 & 0 \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix} \\ y_k &= Cx_k + \begin{bmatrix} 0 & D_2 \end{bmatrix} \begin{bmatrix} w_k \\ v_k \end{bmatrix} \end{aligned}$$

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \bar{D}_1 \bar{w}_k \\ y_k &= Cx_k + \bar{D}_2 \bar{w}_k \end{aligned}$$

Process noise $\neq$ Measurement noise

Optimal State Estimation

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$
$$y_k = Cx_k + D_2 w_k$$

Assumptions

- A, B, C, D_1, D_2 known
- u_k known
- $w_k \sim \mathcal{N}(0, I), \quad \mathbb{E}[x_0] = \bar{x}$
- $\mathrm{Cov}[x_0, w_k] = 0$

Observations

- Since w_k is a RV, x_k and y_k are random vectors
- Since w_k does not affect x_k, $\mathrm{Cov}[x_k, w_k] = 0$

Optimal State Estimator

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$
$$y_k = Cx_k + D_2 w_k$$

- Consider the estimator

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K(y_k - C\hat{x}_k)$$

  - Since y_k is an RV, $\hat{x}_k$ is an RV
  - The state error $e_k \triangleq \hat{x}_k - x_k$ is also an RV with covariance $P_k = \mathrm{Cov}[e_k]$

- Optimal State Estimator chooses K that minimizes the covariance of the state error

$$K_k = \min_{\hat{K} \in \mathbb{R}^{l_x \times l_y}} \text{trace}\, P_{k+1}$$

Optimal State Estimator

$$P_k = \mathrm{Cov}[e_k] = \mathbb{E}[(e_k - \mathbb{E}[e_k])(e_k - \mathbb{E}[e_k])^{\mathrm{T}}]$$

- We would like $\mathbb{E}[e_k] = 0$.
  - The state error satisfies

$$e_{k+1} = (A - KC)e_k + (KD_2 - D_1)w_k$$
$$\mathbb{E}[e_{k+1}] = \mathbb{E}[(A - KC)e_k] = (A - KC)\mathbb{E}[e_k]$$

  - If $\mathbb{E}[e_k] = 0$, then $\mathbb{E}[e_{k+1}] = 0$.
  - If $\mathbb{E}[e_0] = 0$, then, $\mathbb{E}[e_1] = \mathbb{E}[e_2] = \cdots = 0$.

$$\mathbb{E}[e_0] = \mathbb{E}[\hat{x}_0 - x_0] = \mathbb{E}[\hat{x}_0] - \mathbb{E}[x_0] = \mathbb{E}[\hat{x}_0] - \bar{x}$$

- If $\mathbb{E}[\hat{x}_0] = \bar{x}$. Then, $\mathbb{E}[e_k] = 0$.

Optimal State Estimator

- With $\mathbb{E}[e_k] = 0$,

$$P_k = \mathbb{E}[e_k {e_k}^{\mathrm{T}}]$$

$$P_{k+1} = \mathbb{E}\left[e_{k+1} e_{k+1}^{\mathrm{T}}\right]$$

$$= \mathbb{E}\left[\big((A-KC)e_k + (KD_2 - D_1)w_k\big)\big((A-KC)e_k + (KD_2 - D_1)w_k\big)^{\mathrm{T}}\right]$$

$$= \mathbb{E}\big[(A-KC)e_k e_k^{\mathrm{T}}(A-KC)^{\mathrm{T}} + (KD_2 - D_1)w_k w_k^{\mathrm{T}}(KD_2 - D_1)^{\mathrm{T}}\big]$$

$$= (A-KC)P_k(A-KC)^{\mathrm{T}} + (KD_2 - D_1)(KD_2 - D_1)^{\mathrm{T}}$$

$$= AP_kA^{\mathrm{T}} - KCP_kA^{\mathrm{T}} - AP_kC^{\mathrm{T}}K^{\mathrm{T}} + KCP_kC^{\mathrm{T}}K^{\mathrm{T}}$$
$$+KRK^{\mathrm{T}} - SK^{\mathrm{T}} - KS^{\mathrm{T}} + Q$$

where $Q = D_1 D_1^{\mathrm{T}},\ R = D_2 D_2^{\mathrm{T}},\ S = D_1 D_2^{\mathrm{T}}$

Optimal State Estimator

$$
\begin{aligned}
P_{k+1} &= AP_kA^\mathrm{T} - KCP_kA^\mathrm{T} - AP_kC^\mathrm{T}K^\mathrm{T} + KCP_kC^\mathrm{T}K^\mathrm{T} \\
&\qquad +KRK^\mathrm{T} - SK^\mathrm{T} - KS^\mathrm{T} + Q \\
&= AP_kA^\mathrm{T} + Q - K\left(AP_kC^\mathrm{T} + S\right)^\mathrm{T} - \left(AP_kC^\mathrm{T} + S\right)K^\mathrm{T} \\
&\qquad +K\left(CP_kC^\mathrm{T} + R\right)K^\mathrm{T}
\end{aligned}
$$

$$K_k = \min_{\hat{K}\in\mathbb{R}^{l_x \times l_y}} \operatorname{trace} P_{k+1}$$

$$\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}(XA) = A^\mathrm{T}$$

$$\frac{\mathrm{d}}{\mathrm{d}X}\mathrm{tr}\left(XAX^\mathrm{T}\right) = XA^\mathrm{T} + XA$$

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

$$\frac{\mathrm{d}}{\mathrm{d}K}\operatorname{trace} P_{k+1} = -\left(AP_kC^\mathrm{T} + S\right) + K\left(CP_kC^\mathrm{T} + R\right)$$

$$K_k = \left(AP_kC^\mathrm{T} + S\right)\left(CP_kC^\mathrm{T} + R\right)^{-1}$$

Optimal State Estimator

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K_k(y_k - C\hat{x}_k)$$

$$K_k = \left(AP_kC^{\mathrm{T}} + S\right)\left(CP_kC^{\mathrm{T}} + R\right)^{-1}$$

$$P_{k+1} = AP_kA^{\mathrm{T}} + Q - K_k\left(AP_kC^{\mathrm{T}} + S\right)^{\mathrm{T}}$$

Optimal State Estimator - Simple Example

$$x_{k+1} = x_k + u_k + [1 \quad 0]w_k$$
$$y_k = x_k + [0 \quad 1]w_k$$

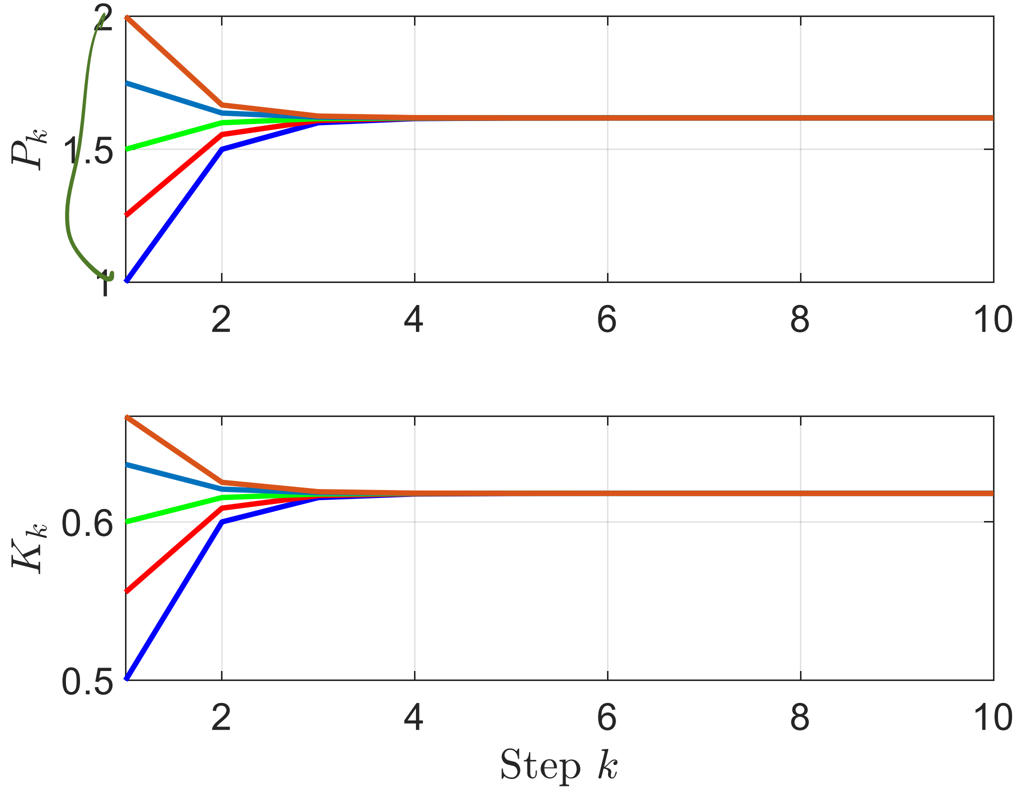
$$\hat{x}_{k+1} = \hat{x}_k + u_k + K_k(y_k - \hat{x}_k)$$

$$K_k = \frac{P_k}{P_k + 1}, \qquad P_{k+1} = \frac{2P_k + 1}{P_k + 1}$$

- If P_k converges, then $P_{k+1} - P_k \to 0$

$$P_\infty = \frac{1}{2}\left(1 + \sqrt{5}\right) \approx 1.62$$

$$K_\infty = \frac{1 + \sqrt{5}}{2 + \sqrt{5}} \approx 0.62$$



Does OSE Improve Accuracy?

$$x_{k+1} = ax_k + bu_k + [\sqrt{q} \quad 0]w_k \text{, } where\ a<1$$
$$y_k = x_k + [0 \quad \sqrt{r}]w_k$$

$$\hat{x}_{k+1} = a\hat{x}_k + bu_k + K_k(y_k - \hat{x}_k)$$
$$K_k = \frac{aP_k}{P_k + r}$$
$$P_{k+1} = a^2 P_k + q - K_k a P_k$$

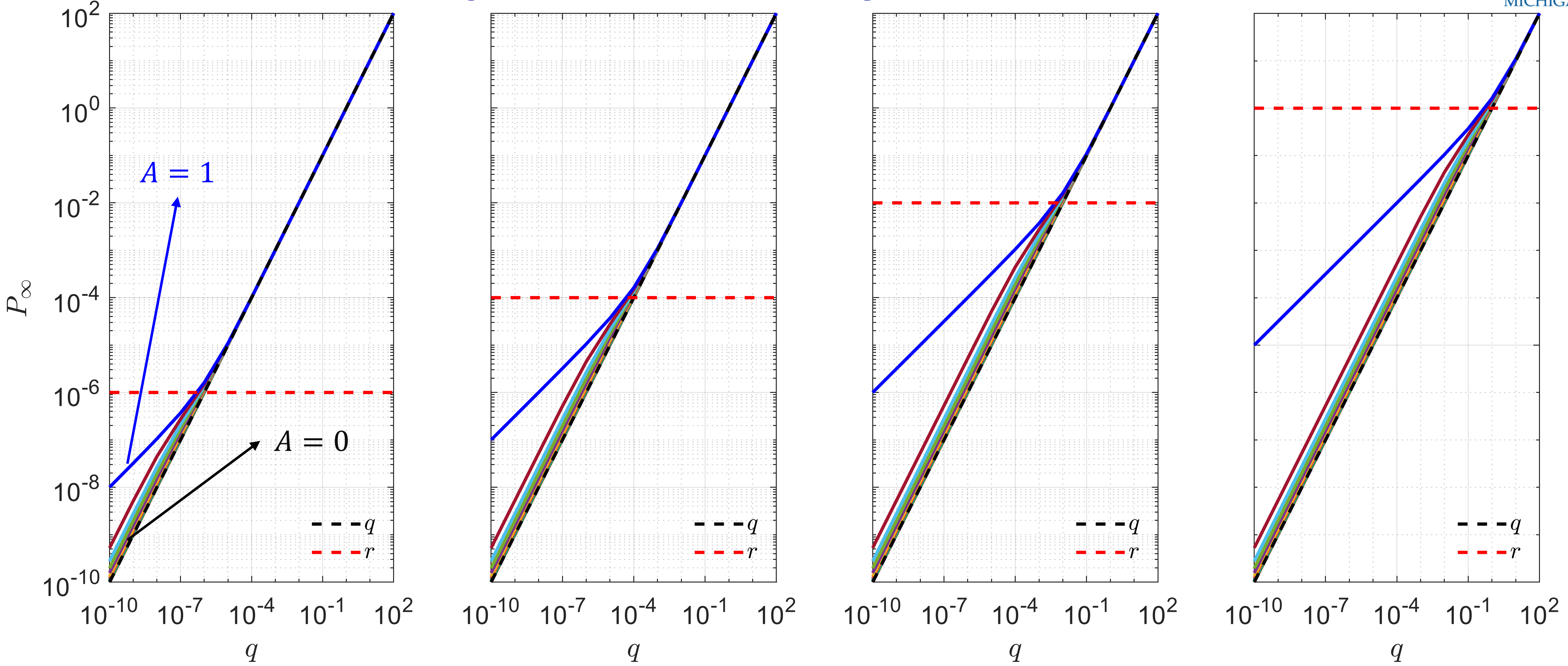
Does OSE Improve Accuracy?

If P_k converges to P_∞ , then $P_\infty = \frac{a^2 r P_\infty}{P_\infty + r} + q$

$$P_\infty = \frac{1}{2}\left(q - (1 - a^2)r + \sqrt{(q - (1 - a^2)r)^2 + 4qr}\right)$$

Case	P_∞	Observation
$q = 0$	0	If the model is good, no matter how bad the sensor is, the estimator state converges to the true state IF $a <= 1$
$r = 0$	q	Even if the sensor is perfect, state estimate accuracy is bounded by the model uncertainty

Does OSE Improve Accuracy?



- If $q > r$, then $P_\infty > q > r$
- If $r > q$, then $r > P_\infty > q$

What if the Noise has a Bias?

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$
$$y_k = Cx_k + D_2 w_k$$

- Let $w_k = \bar{w} + \tilde{w}_k$

- Idea: Estimate $\bar{w}$ with the state!!

- Let $X_k = \begin{bmatrix} x_k \\ \bar{w} \end{bmatrix}$. Then, $X_{k+1} = \begin{bmatrix} x_{k+1} \\ \bar{w} \end{bmatrix} = \begin{bmatrix} Ax_k + Bu_k + D_1\bar{w} + D_1\tilde{w}_k \\ \bar{w} \end{bmatrix}$

$$= \begin{bmatrix} A & D_1 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ \bar{w} \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_k + \begin{bmatrix} D_1 \\ 0 \end{bmatrix} \tilde{w}_k$$

$$= A_\mathrm{a} X_k + B_\mathrm{a} u_k + D_{1\mathrm{a}} \tilde{w}_k$$

What if the Noise has a Bias?

$$x_{k+1} = Ax_k + Bu_k + D_1 w_k$$
$$y_k = Cx_k + D_2 w_k$$

$$y_k = Cx_k + D_2\bar{w} + D_2\widetilde{w}_k$$
$$= \begin{bmatrix} C & D_2 \end{bmatrix} \begin{bmatrix} x_k \\ \bar{w} \end{bmatrix} + D_2\widetilde{w}_k$$
$$= C_{\mathrm{a}}X_k + D_2\tilde{v}_k$$

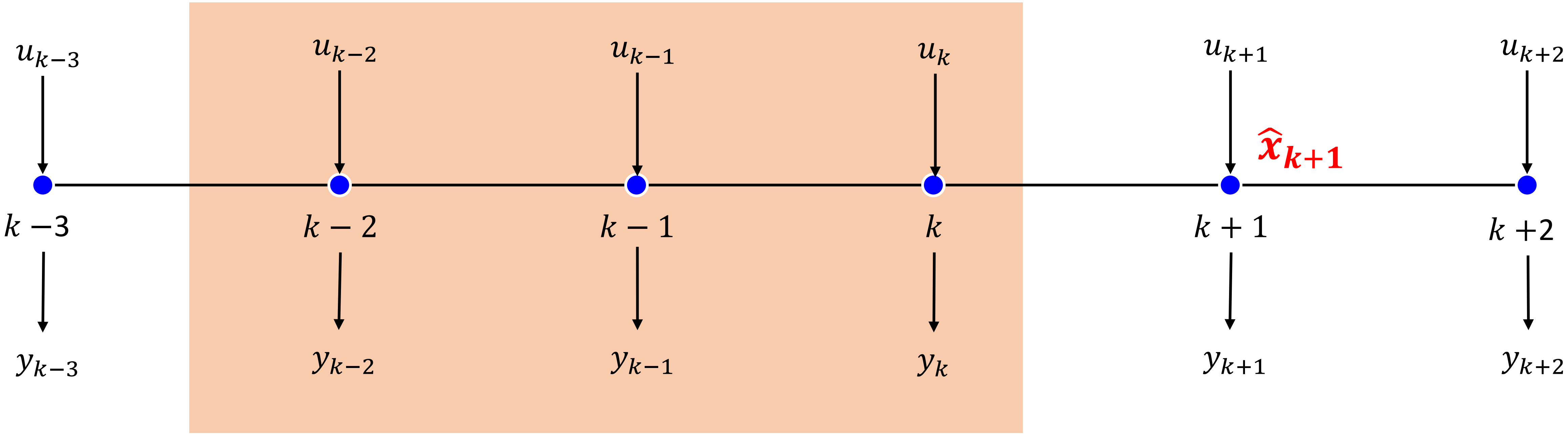
$$X_{k+1} = A_{\mathrm{a}}X_k + B_{\mathrm{a}}u_k + D_{1\mathrm{a}}\widetilde{w}_k$$
$$y_k = C_{\mathrm{a}}X_k + D_2\widetilde{w}_k$$

- This is how we will deal with the gyro bias!!

State Estimators

Predictors

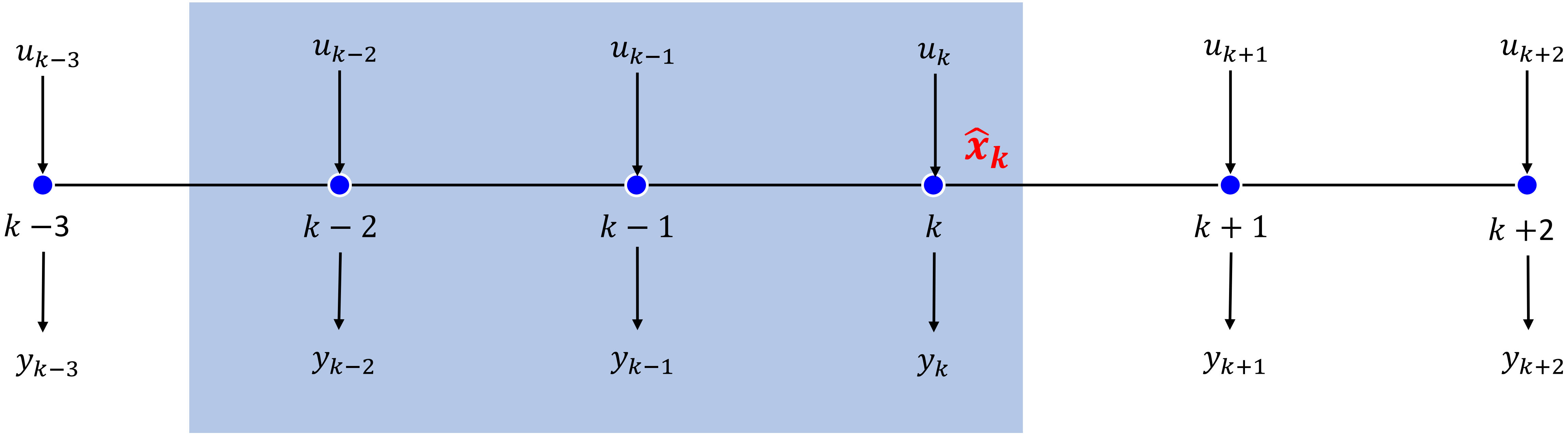
Filters



State Estimators

Predictors

Filters



Summary

- Using measurements, state estimators provide an estimate of the unmeasured states

- In the next lecture,
  - We will derive the equations for the Kalman Filter