State Estimation

Robert Stengel
Robotics and Intelligent Systems MAE 345, Princeton University, 2017

Learning Objectives

- **Compute least-squares estimates of a constant vector**
  - Unweighted and weighted batch processing of noisy data
  - Recursive processing to incorporate new data
- **Estimate the state of an uncertain linear dynamic system with incomplete, noisy measurements**
  - Discrete-time Kalman filter
  - Continuous-time Kalman-Bucy filter (supplement)

Estimate Constant Vector by Inverse Transformation

- **Given**
  - Measurements, **y**, of a constant vector, **x**
- **Estimate x**
- **Assume that output, y, is a perfect measurement and H is invertible**

$$\boxed{\mathbf{y} = \mathbf{H}\,\mathbf{x}}$$

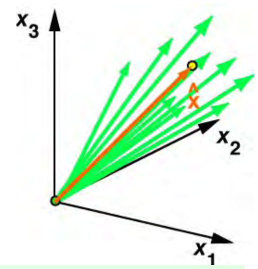
- **y**: (n x 1) output vector
- **H**: (n x n) output matrix
- **x** : (n x 1) vector to be estimated

- **Estimate is based on inverse transformation**

$$\boxed{\hat{\mathbf{x}} = \mathbf{H}^{-1}\,\mathbf{y}}$$

Imperfect Measurement of a Constant Vector

- Given
  - "Noisy" measurements, **z**, of a constant vector, **x**
- Effects of error can be reduced if measurement is redundant
- Noise-free output, **y**



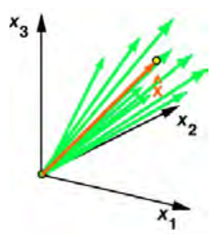
$$\mathbf{y} = \mathbf{H}\,\mathbf{x}$$

- **y**: (k x 1) output vector
- **H**: (k x n) output matrix, $k > n$
- **x** : (n x 1) vector to be estimated

- Measurement of output with error, **z**

$$\mathbf{z} = \mathbf{y} + \mathbf{n} = \mathbf{H}\,\mathbf{x} + \mathbf{n}$$

- **z**: (k x 1) measurement vector
- **n** : (k x 1) error vector



Least-Squares Estimate of a Constant Vector

- Measurement-error residual

$$\boldsymbol{\varepsilon} = \mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}} = \mathbf{z} - \hat{\mathbf{y}}$$

$$\dim(\boldsymbol{\varepsilon}) = (k \times 1)$$

- Squared measurement error = cost function, J

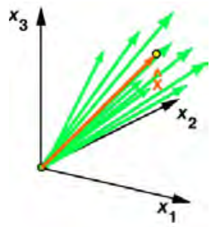
$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \frac{1}{2}(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}})^T(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}})$$
$$= \frac{1}{2}\left(\mathbf{z}^T\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{z} - \mathbf{z}^T\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)$$

Quadratic norm

- What is the control parameter?

The estimate of x $\hat{\mathbf{x}}$

$$\dim(\hat{\mathbf{x}}) = (n \times 1)$$



Least-Squares Estimate of a Constant Vector

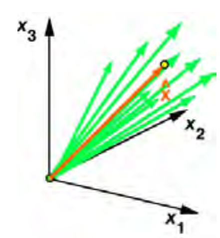
$$J = \frac{1}{2}\left(\mathbf{z}^T\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{z} - \mathbf{z}^T\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)$$

Necessary condition for a minimum

$$\begin{aligned}\frac{\partial J}{\partial \hat{\mathbf{x}}} &= \mathbf{0} \\ &= \frac{1}{2}\left[\mathbf{0} - \left(\mathbf{H}^T\mathbf{z}\right)^T - \mathbf{z}^T\mathbf{H} + \left(\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)^T + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\right]\end{aligned}$$

The 2nd and 4th terms are transposes of the 3rd and 5th terms

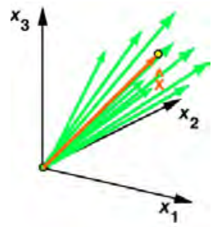
Least-Squares Estimate of a Constant Vector



The derivative of a scalar, *J*, with respect to a vector, x, (i.e., the gradient) is defined to be a row vector; thus,

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \left[\begin{array}{cccc} \dfrac{\partial J}{\partial \hat{x}_1} & \dfrac{\partial J}{\partial \hat{x}_2} & \cdots & \dfrac{\partial J}{\partial \hat{x}_n} \end{array}\right]$$

$$\begin{aligned}\mathbf{0} &= \frac{1}{2}\left[\mathbf{0} - \left(\mathbf{H}^T\mathbf{z}\right)^T - \mathbf{z}^T\mathbf{H} + \left(\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)^T + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\right] \\ &= \left[-\mathbf{z}^T\mathbf{H} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\right]\end{aligned}$$



Optimal Estimate of x

Rearranging

$$\left[\hat{\mathbf{x}}^T \mathbf{H}^T \mathbf{H} - \mathbf{z}^T \mathbf{H}\right] = \mathbf{0}$$

$$\hat{\mathbf{x}}^T \mathbf{H}^T \mathbf{H} = \mathbf{z}^T \mathbf{H}$$

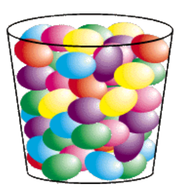
Then, the optimal estimate is

$$\left[\hat{\mathbf{x}}^T \left(\mathbf{H}^T \mathbf{H}\right)\right]\left(\mathbf{H}^T \mathbf{H}\right)^{-1} = \hat{\mathbf{x}}^T \left(\mathbf{I}\right)$$

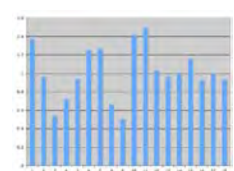
$$\hat{\mathbf{x}}^T = \mathbf{z}^T \mathbf{H}\left(\mathbf{H}^T \mathbf{H}\right)^{-1} \quad (row)$$

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T \mathbf{H}\right)^{-1} \mathbf{H}^T \mathbf{z} \quad (column)$$

$(H^T H)^{-1} H^T$ is called the *left pseudoinverse* of H



Estimate of Scalar Constant: Average Weight of Jelly Beans



Measurements are equally uncertain

$$z_i = x + n_i \quad ,i = 1 \; to \; k$$

Express measurements as

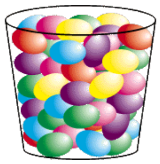
$$\mathbf{z} = \mathbf{H}x + \mathbf{n}$$

Output matrix

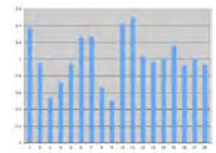
$$\mathbf{H} = \begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}$$

Optimal estimate

$$\hat{x} = \left(\mathbf{H}^T \mathbf{H}\right)^{-1} \mathbf{H}^T \mathbf{z}$$



Estimation of a Scalar Constant: Average Weight of the Jelly Beans



Optimal estimate

$$\hat{x} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\,\mathbf{z}$$

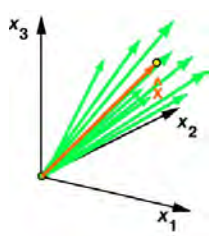
$$(1\times 1) = \left[(1\times k)(k\times 1)\right]^{-1}(1\times k)(k\times 1)$$

$$\hat{x} = \left(\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ \ldots \\ z_k \end{bmatrix}$$

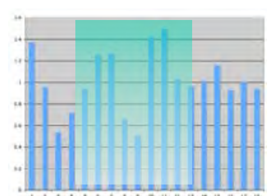
...is the average

$$\hat{x} = (k)^{-1}\left(z_1 + z_2 + \ldots + z_k\right)$$

$$\hat{x} = \frac{1}{k}\sum_{i=1}^{k} z_i \quad \left[\textit{sample mean value}\right]$$



Measurements of Differing Quality



Original cost function, *J*, and optimal estimate of x

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z}-\mathbf{H}\,\hat{\mathbf{x}}\right)^T\left(\mathbf{z}-\mathbf{H}\,\hat{\mathbf{x}}\right)$$
$$= \frac{1}{2}\left(\mathbf{z}^T\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{z} - \mathbf{z}^T\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{H}\,\hat{\mathbf{x}}\right)$$

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\,\mathbf{z}$$

Suppose some elements of measurement, z, are more uncertain than others

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n}$$

$\mathbf{n}$: Error vector

Give more uncertain measurements less weight in arriving at minimum-cost estimate

Error-Weighted Cost Function

Measurement uncertainty matrix, R (large is worse)

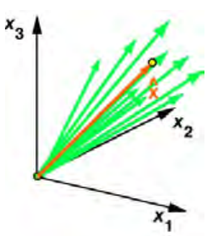
$$\mathbf{R} = \begin{bmatrix} \text{(large error)} & 0 & \ldots & 0 \\ 0 & \text{(small error)} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \text{(medium error)} \end{bmatrix}$$

Error-weighting matrix, $\mathbf{R}^{-1}$

$$\mathbf{R}^{-1} = \begin{bmatrix} \text{(low weight)} & 0 & \ldots & 0 \\ 0 & \text{(high weight)} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \text{(medium weight)} \end{bmatrix}$$

Weighted cost function, *J*, reduces significance of poorer measurements

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T \mathbf{R}^{-1}\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\right)^T \mathbf{R}^{-1}\left(\mathbf{z} - \mathbf{H}\hat{\mathbf{x}}\right)$$
$$= \frac{1}{2}\left(\mathbf{z}^T\mathbf{R}^{-1}\mathbf{z} - \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z} - \mathbf{z}^T\mathbf{R}^{-1}\mathbf{H}\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\hat{\mathbf{x}}\right)$$



Weighted Least-Squares Estimate of a Constant Vector

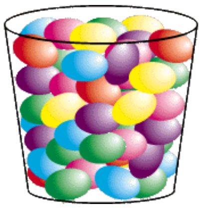
Necessary condition for minimum

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \mathbf{0} = \frac{1}{2}\left[\mathbf{0} - \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}\right)^T - \mathbf{z}^T\mathbf{R}^{-1}\mathbf{H} + \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\hat{\mathbf{x}}\right)^T + \hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]$$

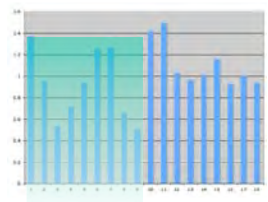
$$\left[\hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} - \mathbf{z}^T\mathbf{R}^{-1}\mathbf{H}\right] = \mathbf{0}$$
$$\hat{\mathbf{x}}^T\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} = \mathbf{z}^T\mathbf{R}^{-1}\mathbf{H}$$

Weighted optimal estimate:

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}$$



Weighted Estimate of Average Jelly Bean Weight



Error-weighting matrix based on standard deviations

$$\mathbf{R}^{-1} = \mathbf{A} = \begin{bmatrix} 1/\sigma_{n_1}^2 & 0 & \ldots & 0 \\ 0 & 1/\sigma_{n_2}^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1/\sigma_{n_k}^2 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{kk} \end{bmatrix}$$

Optimal estimate of average jelly bean weight

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{z}$$

$$\hat{x} = \left(\begin{bmatrix} 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & a_{kk} \end{bmatrix} \begin{bmatrix} 1 \\ \ldots \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & \ldots & 1 \end{bmatrix} \begin{bmatrix} a_{11} & \ldots & 0 \\ \ldots & \ldots & \ldots \\ 0 & \ldots & a_{kk} \end{bmatrix} \begin{bmatrix} z_1 \\ \ldots \\ z_k \end{bmatrix}$$

$$\hat{x} = \frac{\sum_{i=1}^{k} a_{ii} z_i}{\sum_{i=1}^{k} a_{ii}}$$

Recursive Least-Squares Estimate of Constant Vector, x

- "Batch-processing" approach
  - All information is gathered prior to processing
  - All information is processed at once

- Recursive approach
  - Optimal estimate has been made from prior measurement set
  - Additional new measurement set
  - Optimal estimate improved by correction to prior estimate

Add One New Measurement

Initial measurement set and state estimate

$$\mathbf{z}_1 = \mathbf{H}_1\mathbf{x} + \mathbf{n}_1, \quad \dim(\mathbf{z}_1) = k_1 \times 1$$

$$\hat{\mathbf{x}}_1 = \left(\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{H}_1\right)^{-1}\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{z}_1$$

$\mathbf{R}_1$: Error covariance of 1st measurement

New measurement set

$$\mathbf{z}_2 = \mathbf{H}_2\mathbf{x} + \mathbf{n}_2, \quad \dim(\mathbf{z}_2) = k_2 \times 1$$

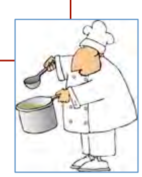
$\mathbf{R}_2$: Error covariance of 2nd measurement

Improved Estimate Incorporating New Measurement Set

$$\mathbf{P}_1^{-1} \triangleq \mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{H}_1 \qquad \hat{\mathbf{x}}_1 = \mathbf{P}_1\mathbf{H}_1^T\mathbf{R}_1^{-1}\mathbf{z}_1$$

New estimate is linear correction to old

$$\begin{aligned}\hat{\mathbf{x}}_2 &= \hat{\mathbf{x}}_1 - \mathbf{P}_1\mathbf{H}_2^T\left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}\left(\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_1\right) \\ &\triangleq \hat{\mathbf{x}}_1 - \mathbf{K}\left(\mathbf{z}_2 - \mathbf{H}_2\hat{\mathbf{x}}_1\right)\end{aligned}$$



$\mathbf{K}$: Estimator gain matrix

$$= \mathbf{P}_1\mathbf{H}_2^T\left(\mathbf{H}_2\mathbf{P}_1\mathbf{H}_2^T + \mathbf{R}_2\right)^{-1}$$

See reading for details

Recursive Optimal Estimate of Constant Vector, x

- Sequence of new measurements
- Generalize to a recursive form, with index *i*

$$\boxed{\hat{\mathbf{x}}_i = \hat{\mathbf{x}}_{i-1} - \mathbf{K}_i\left(\mathbf{z}_i - \mathbf{H}_i\hat{\mathbf{x}}_{i-1}\right)}$$

with

$$\mathbf{K}_i = \mathbf{P}_{i-1}\mathbf{H}_i^T\left(\mathbf{H}_i\mathbf{P}_{i-1}\mathbf{H}_i^T + \mathbf{R}_i\right)^{-1}$$

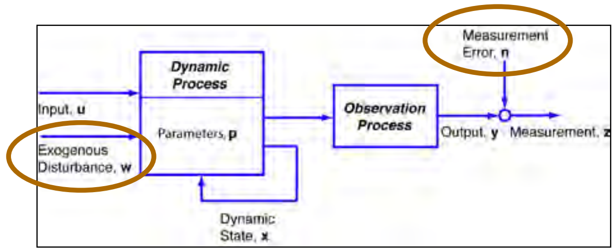
$$\mathbf{P}_i = \left(\mathbf{P}_{i-1}^{-1} + \mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\right)^{-1}$$

$\dim(\mathbf{x}) = n \times 1;\quad \dim(\mathbf{P}) = n \times n$
$\dim(\mathbf{z}) = r \times 1;\quad \dim(\mathbf{R}) = r \times r$
$\dim(\mathbf{H}) = r \times n;\quad \dim(\mathbf{K}) = n \times r$

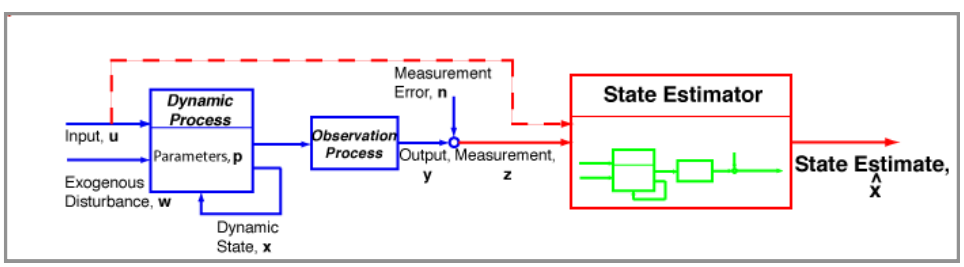
Dynamic Sampled-Data Systems with Uncertain Inputs and Disturbances

Systems with Uncertainty



- **x is not constant in a dynamic system**
- **Dynamic systems may have uncertain**
  - Initial conditions
  - Inputs
  - Measurements
  - System parameters or dynamic structure
- **Design goal: estimate the state with minimum expected error**
  - Mean value ➔ actual value of the state
  - Expected value of estimate error as small as possible

State Estimation



- **Goals**
  - Minimize effects of <u>measurement error</u> on knowledge of state
  - Recontruct full state from <u>reduced measurement set</u> ($r \le n$)
  - Average <u>redundant measurements</u> ($r \ge n$) to produce estimate of full state
- **Method**
  - Provide optimal balance between measurements and estimates based on dynamic model alone
  - Continuous- or discrete-time implementation

Uncertain Sampled-Data Linear Dynamic Model

Discrete-time LTI model with known coefficients

$$\boxed{\mathbf{x}_k = \mathbf{\Phi x}_{k-1} + \mathbf{\Gamma u}_{k-1} + \mathbf{\Lambda w}_{k-1}}$$

$$\boxed{\begin{aligned}\mathbf{y}_k &= \mathbf{H}_{\mathbf{x}}\mathbf{x}_k + \mathbf{H}_u\mathbf{u}_k \\ \mathbf{z}_k &= \mathbf{y}_k + \mathbf{n}_k\end{aligned}}$$

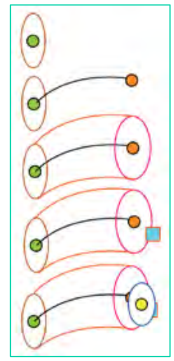
Equivalent to continuous-time model at sampling instants

Optimal State Estimation

see Supplemental Material for continuous-time filter and example

Discrete-Time Linear-Optimal State Estimation

- Kalman filter is optimal estimator for discrete-time linear systems with Gaussian uncertainty
- Five equations
  1) State estimate extrapolation
  2) Covariance estimate extrapolation
  3) Filter gain computation
  4) State estimate update
  5) Covariance estimate "update"



- Notation

$$\boxed{\begin{array}{ll} \hat{\mathbf{x}}_k(-): & \text{Estimate at } k^{th} \text{ instant } \textbf{before} \text{ measurement update} \\ \hat{\mathbf{x}}_k(+): & \text{Estimate at } k^{th} \text{ instant } \textbf{after} \text{ measurement update} \end{array}}$$

Equations of the Kalman Filter

1) State estimate extrapolation (or propagation)

$$\boxed{\hat{\mathbf{x}}_k(-) = \mathbf{\Phi}_{k-1}\,\hat{\mathbf{x}}_{k-1}(+) + \mathbf{\Gamma}_{k-1}\mathbf{u}_{k-1}}$$

2) Covariance estimate extrapolation (or propagation)

$$\boxed{\mathbf{P}_k(-) = \mathbf{\Phi}_{k-1}\,\mathbf{P}_{k-1}(+)\mathbf{\Phi}_{k-1}^T + \mathbf{Q}_{k-1}}$$

Equations of the Kalman Filter

3) Filter gain computation

$$\boxed{\mathbf{K}_k = \mathbf{P}_k(-)\mathbf{H}_k^T\left[\mathbf{H}_k\mathbf{P}_k(-)\mathbf{H}_k^T + \mathbf{R}_k\right]^{-1}}$$

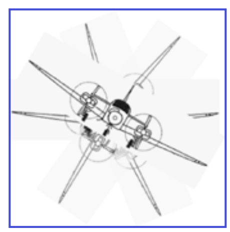
4) State estimate update

$$\boxed{\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + \mathbf{K}_k\left[\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_k(-)\right]}$$

5) Covariance estimate "update"

$$\boxed{\mathbf{P}_k(+) = \left[\mathbf{P}_k^{-1}(-) + \mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\right]^{-1}}$$

<u>Example</u>: Estimate Rolling Motion of an Airplane



<u>Continuous-time</u> model

$$\begin{bmatrix}\Delta\dot{p}\\ \Delta\dot{\phi}\end{bmatrix} = \begin{bmatrix}L_p & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}\Delta p\\ \Delta\phi\end{bmatrix} + \begin{bmatrix}L_{\delta A}\\ 0\end{bmatrix}\Delta\delta A + \begin{bmatrix}L_p\\ 0\end{bmatrix}\Delta p_w$$

$$\begin{bmatrix}\Delta p\\ \Delta\phi\end{bmatrix} = \begin{bmatrix}\text{Roll rate, rad/s}\\ \text{Roll angle, rad}\end{bmatrix}$$

$\Delta\delta A$ = Aileron deflection, rad

Δp_w = Turbulence disturbance, rad/s

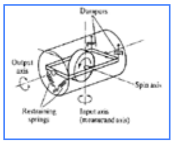
<u>Discrete-time</u> model

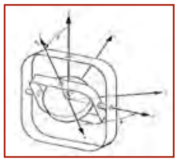
$$\begin{bmatrix}\Delta p_k\\ \Delta\phi_k\end{bmatrix} = \begin{bmatrix}e^{L_pT} & 0\\ \dfrac{\left(e^{L_pT}-1\right)}{L_p} & 1\end{bmatrix}\begin{bmatrix}\Delta p_{k-1}\\ \Delta\phi_{k-1}\end{bmatrix} + \begin{bmatrix}\sim L_{\delta A}T\\ 0\end{bmatrix}\Delta\delta A_{k-1} + \begin{bmatrix}\sim L_pT\\ 0\end{bmatrix}\Delta p_{w_{k-1}}$$

$$= \begin{bmatrix}\varphi_{11} & 0\\ \varphi_{21} & 1\end{bmatrix}\begin{bmatrix}\Delta p_{k-1}\\ \Delta\phi_{k-1}\end{bmatrix} + \begin{bmatrix}\gamma_1\\ 0\end{bmatrix}\Delta\delta A_{k-1} + \begin{bmatrix}\lambda_1\\ 0\end{bmatrix}\Delta p_{w_{k-1}}$$

T = sampling interval, s

Kalman Filter Example





Rate and Angle Measurement

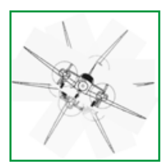
$$\begin{matrix}\text{Roll rate gyro, rad/s}\\ \text{Roll angle gyro, rad}\end{matrix}\begin{bmatrix}\Delta p_M\\ \Delta\phi_M\end{bmatrix}_k=\begin{bmatrix}\Delta p+\Delta n_p\\ \Delta\phi+\Delta n_\phi\end{bmatrix}_k$$

$$=\mathbf{I}\Delta\mathbf{x}_k+\Delta\mathbf{n}_k$$

1) State Estimate Extrapolation

$$\begin{bmatrix}\Delta\hat{p}_k(-)\\ \Delta\hat{\phi}_k(-)\end{bmatrix}=\begin{bmatrix}\varphi_{11} & 0\\ \varphi_{21} & 1\end{bmatrix}\begin{bmatrix}\Delta\hat{p}_{k-1}(+)\\ \Delta\hat{\phi}_{k-1}(+)\end{bmatrix}+\begin{bmatrix}\gamma_1\\ 0\end{bmatrix}\Delta\delta A_{k-1}$$

Kalman Filter Example



2) Covariance Extrapolation

$$\begin{bmatrix}p_{11}(-) & p_{12}(-)\\ p_{21}(-) & p_{22}(-)\end{bmatrix}_k=\begin{bmatrix}\varphi_{11} & 0\\ \varphi_{21} & 1\end{bmatrix}\begin{bmatrix}p_{11}(+) & p_{12}(+)\\ p_{21}(+) & p_{22}(+)\end{bmatrix}_{k-1}\begin{bmatrix}\varphi_{11} & \varphi_{21}\\ 0 & 1\end{bmatrix}+\begin{bmatrix}\sigma_p^2 & 0\\ 0 & 0\end{bmatrix}$$

where

$$\mathbf{Q}_{k-1}\approx\begin{bmatrix}L_p\\ 0\end{bmatrix}Q'_C\begin{bmatrix}L_p & 0\end{bmatrix}T=\begin{bmatrix}\sigma_p^2 & 0\\ 0 & 0\end{bmatrix}$$

3) Gain Computation

$$\begin{bmatrix}k_{11} & k_{12}\\ k_{21} & k_{22}\end{bmatrix}_k=\begin{bmatrix}p_{11}(-) & p_{12}(-)\\ p_{21}(-) & p_{22}(-)\end{bmatrix}_k\left\{\begin{bmatrix}p_{11}(-) & p_{12}(-)\\ p_{21}(-) & p_{22}(-)\end{bmatrix}_k+\begin{bmatrix}\sigma_{p_M}^2 & 0\\ 0 & \sigma_{\phi_M}^2\end{bmatrix}_k\right\}^{-1}$$

where

$$\mathbf{R}_k\delta_{jk}=\begin{bmatrix}\sigma_{p_M}^2 & 0\\ 0 & \sigma_{\phi_M}^2\end{bmatrix}_k$$

Kalman Filter Example

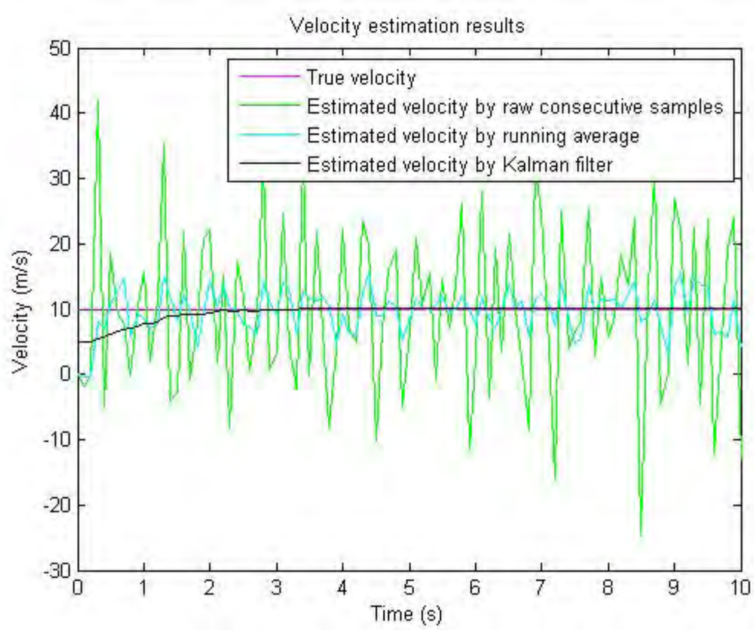
4) State Estimate Update

$$\begin{bmatrix} \Delta\hat{p}_k(+) \\ \Delta\hat{\phi}_k(+) \end{bmatrix} = \begin{bmatrix} \Delta\hat{p}_k(-) \\ \Delta\hat{\phi}_k(-) \end{bmatrix} + \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}_k \left\{ \begin{bmatrix} \Delta p_{M_k} \\ \Delta \phi_{M_k} \end{bmatrix} - \begin{bmatrix} \Delta\hat{p}_k(-) \\ \Delta\hat{\phi}_k(-) \end{bmatrix} \right\}$$

5) Covariance "Update"

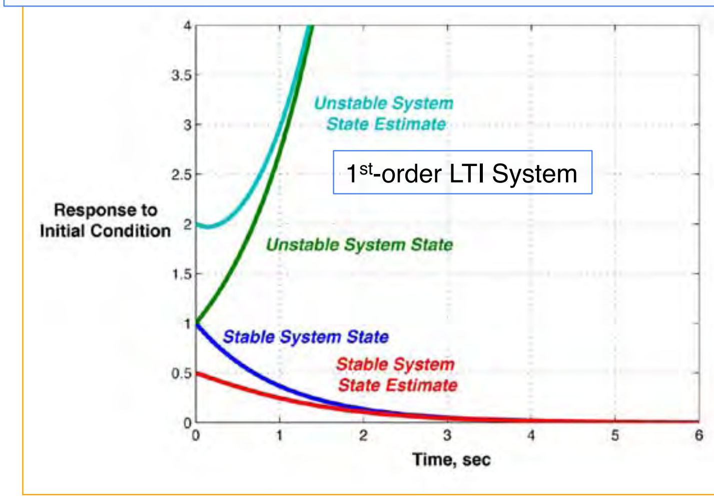
$$\begin{bmatrix} p_{11}(+) & p_{12}(+) \\ p_{21}(+) & p_{22}(+) \end{bmatrix}_k = \left\{ \begin{bmatrix} p_{11}(-) & p_{12}(-) \\ p_{21}(-) & p_{22}(-) \end{bmatrix}_k + \begin{bmatrix} \frac{1}{\sigma_{p_M}^2} & 0 \\ 0 & \frac{1}{\sigma_{\phi_M}^2} \end{bmatrix}_k \right\}^{-1}$$

Comparison of Running Average and Kalman Estimate of Velocity from Position Measurement



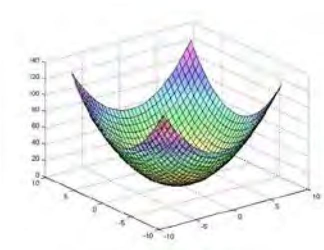
Kalman Filter Estimate is Stable

Estimate is stable even if dynamic system is unstable, i.e., asymptotically approaches actual state



Next Time: Stochastic Control

Supplemental Material



Is the Least-Squares Solution a Minimum or a Maximum?

Gradient

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \begin{bmatrix} \dfrac{\partial J}{\partial \hat{x}_1} & \dfrac{\partial J}{\partial \hat{x}_2} & \cdots & \dfrac{\partial J}{\partial \hat{x}_n} \end{bmatrix} = \begin{bmatrix} -\mathbf{z}^T \mathbf{H} + \hat{\mathbf{x}}^T \mathbf{H}^T \mathbf{H} \end{bmatrix} = \mathbf{0}$$

Hessian matrix

$$\frac{\partial^2 J}{\partial \hat{\mathbf{x}}^2} = \mathbf{H}^T \mathbf{H} > \mathbf{0}, \quad \dim = (n \times n)$$

A minimum

Weighted Least-Squares Estimate of a Constant Vector

Weighted cost function, *J*

$$J = \frac{1}{2}\boldsymbol{\varepsilon}^T \mathbf{R}^{-1}\boldsymbol{\varepsilon} = \frac{1}{2}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)^T \mathbf{R}^{-1}\left(\mathbf{z} - \mathbf{H}\,\hat{\mathbf{x}}\right)$$
$$= \frac{1}{2}\left(\mathbf{z}^T \mathbf{R}^{-1}\mathbf{z} - \hat{\mathbf{x}}^T \mathbf{H}^T \mathbf{R}^{-1}\mathbf{z} - \mathbf{z}^T \mathbf{R}^{-1}\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^T \mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}\,\hat{\mathbf{x}}\right)$$

Necessary condition for a minimum

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \mathbf{0} = \frac{1}{2}\left[\mathbf{0} - \left(\mathbf{H}^T \mathbf{R}^{-1}\mathbf{z}\right)^T - \mathbf{z}^T \mathbf{R}^{-1}\mathbf{H} + \left(\mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}\,\hat{\mathbf{x}}\right)^T + \hat{\mathbf{x}}^T \mathbf{H}^T \mathbf{R}^{-1}\mathbf{H}\right]$$

Uncertain Continuous-Time Linear Dynamic Model

Continuous-time LTI model with known coefficients

$$\dot{\mathbf{x}}(t) = \mathbf{F}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}\mathbf{w}(t), \quad \mathbf{x}(t_o)\ given$$
$$\mathbf{x}(t) = \mathbf{x}(t_o) + \int_{t_o}^{t}\left[\mathbf{F}\mathbf{x}(\tau) + \mathbf{G}\mathbf{u}(\tau) + \mathbf{L}\mathbf{w}(\tau)\right]d\tau$$

$$\mathbf{y}(t) = \mathbf{H}_{\mathbf{x}}\mathbf{x}(t) + \mathbf{H}_{\mathbf{u}}\mathbf{u}(t): \text{ Output vector}$$
$$\mathbf{z}(t) = \mathbf{y}(t) + \mathbf{n}(t): \text{ Measurement vector}$$

Initial condition and disturbance inputs are not known precisely

Measurement of state is transformed and is subject to error

Continuous-Time Linear-Optimal State Estimation

Continuous-time linear dynamic process with random disturbance

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

Measurement with random error

$$\boxed{\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{n}(t)}$$

Uncertainty model for initial condition, disturbance input, and measurement error

$$\begin{aligned}
\bar{\mathbf{x}}(t_0) &= E\left[\mathbf{x}(t_0)\right] \\
\mathbf{P}(t_0) &= E\left\{\left[\mathbf{x}(t_0) - \bar{\mathbf{x}}(t_0)\right]\left[\mathbf{x}(t_0) - \bar{\mathbf{x}}(t_0)\right]^T\right\} \\
\mathbf{u}(t) &= E\left[\mathbf{u}(t)\right] \\
\mathbf{U}(t_0) &= \mathbf{0}
\end{aligned}$$

$$\begin{aligned}
\bar{\mathbf{w}}(t) &= \mathbf{0} \\
\mathbf{W}(t) &= E\left\{\left[\mathbf{w}(t)\right]\left[\mathbf{w}(\tau)\right]^T\right\} \\
\bar{\mathbf{n}}(t) &= \mathbf{0} \\
\mathbf{N}(t) &= E\left\{\left[\mathbf{n}(t)\right]\left[\mathbf{n}(\tau)\right]^T\right\}
\end{aligned}$$

Linear-Optimal State Estimator (Kalman-Bucy Filter)

- **Optimal estimate of state**

$$\boxed{\begin{aligned}
&\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\hat{\mathbf{x}}(t) + \mathbf{G}(t)\mathbf{u}(t) + \mathbf{K}(t)\left[\mathbf{z}(t) - \mathbf{H}\hat{\mathbf{x}}(t)\right], \quad \hat{\mathbf{x}}(t_o) = \bar{\mathbf{x}}(t_o) \\
&\mathbf{K}(t): \text{ Optimal estimator gain matrix } (n \times r)
\end{aligned}}$$

- **Two parts to the optimal state estimator**
  - **Propagation of the expected value of x**
  - **Least-squares correction to the model-based estimate**

$$\boxed{\Delta\dot{\hat{\mathbf{x}}}(t) = \mathbf{F}(t)\Delta\hat{\mathbf{x}}(t) + \mathbf{G}\Delta\mathbf{u}(t)} \quad \boxed{+\mathbf{K}(t)\left[\Delta\mathbf{z}(t) - \mathbf{H}\Delta\hat{\mathbf{x}}(t)\right]}$$

LTI System $$\boxed{\Delta\dot{\hat{\mathbf{x}}}(t) = \left[\mathbf{F} - \mathbf{K}\mathbf{H}\right]\Delta\hat{\mathbf{x}}(t) + \mathbf{G}\Delta\mathbf{u}(t) + \mathbf{K}\Delta\mathbf{z}(t)}$$

Estimator Gain for the Kalman-Bucy Filter

Optimal filter gain matrix

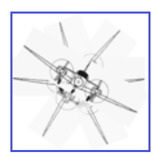
$$\boxed{\mathbf{K}(t)=\mathbf{P}(t)\mathbf{H}^T\mathbf{N}^{-1}(t)}$$

Matrix Riccati equation for estimator covariance

$$\dot{\mathbf{P}}(t)=\mathbf{F}(t)\mathbf{P}(t)+\mathbf{P}(t)\mathbf{F}^T(t)+\mathbf{L}(t)\mathbf{W}(t)\mathbf{L}^T(t)$$
$$-\mathbf{P}(t)\mathbf{H}^T\mathbf{N}^{-1}\mathbf{H}\mathbf{P}(t),\quad \mathbf{P}(t_o)=\mathbf{P}_o$$

- Same equations as those that define LQ control gain, except
  - Solution matrix, **P**, propagated forward in time
  - Matrices are modified

Continuous-Time 2nd-Order Example of Kalman-Bucy Filter



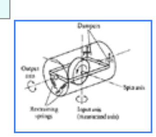
Rolling motion of an airplane

$$\begin{bmatrix}\dot{p}(t)\\ \dot{\phi}(t)\end{bmatrix}=\begin{bmatrix}L_p & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}p(t)\\ \phi(t)\end{bmatrix}+\begin{bmatrix}L_{\delta A}\\ 0\end{bmatrix}\delta A(t)+\begin{bmatrix}L_p\\ 0\end{bmatrix}p_w(t)$$

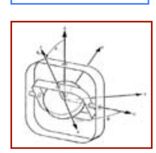
$$\begin{bmatrix}p\\ \phi\end{bmatrix}=\begin{bmatrix}\text{Roll rate, rad/s}\\ \text{Roll angle, rad}\end{bmatrix}$$
$$\delta A=\text{Aileron deflection, rad}$$
$$p_w=\text{Turbulence disturbance, rad/s}$$

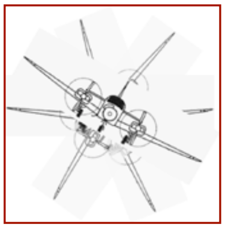
$$L_p:\begin{cases}\textit{Roll-rate damping}\\ \textit{Turbulence sensitivity}\end{cases}$$
$$L_{\delta A}:\ \textit{Control effectiveness}$$

Measurement of roll rate and angle



$$\begin{bmatrix}p_M(t)\\ \phi_M(t)\end{bmatrix}_k=\begin{bmatrix}p(t)+n_p(t)\\ \phi(t)+n_\phi(t)\end{bmatrix}_k=\mathbf{I}_2\mathbf{x}(t)+\mathbf{n}(t)$$





Second-Order Example of Kalman-Bucy Filter

Covariance extrapolation

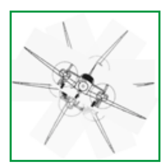
$$\begin{bmatrix} \dot{p}_{11}(t) & \dot{p}_{12}(t) \\ \dot{p}_{12}(t) & \dot{p}_{22}(t) \end{bmatrix} = \begin{bmatrix} L_p & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} + \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} \begin{bmatrix} L_p & 1 \\ 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} L_p^2 \sigma_{p_W}^2 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} \begin{bmatrix} \sigma_{p_M}^2 & 0 \\ 0 & \sigma_{\phi_M}^2 \end{bmatrix}^{-1} \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix}$$

Estimator gain computation

$$\begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{21}(t) & k_{22}(t) \end{bmatrix} = \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} \begin{bmatrix} \sigma_{p_M}^2 & 0 \\ 0 & \sigma_{\phi_M}^2 \end{bmatrix}^{-1}$$

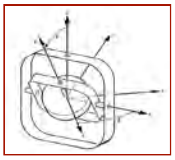
Kalman-Bucy Filter with Two Measurements



State estimate with roll rate and angle measurements

$$\begin{bmatrix} \dot{\hat{p}}(t) \\ \dot{\hat{\phi}}(t) \end{bmatrix} = \begin{bmatrix} L_p & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{p}(t) \\ \hat{\phi}(t) \end{bmatrix} + \begin{bmatrix} L_{\delta A} \\ 0 \end{bmatrix} \delta A(t)$$

$$+ \begin{bmatrix} k_{11}(t) & k_{12}(t) \\ k_{21}(t) & k_{22}(t) \end{bmatrix} \begin{bmatrix} p_M(t) - \hat{p}(t) \\ \phi_M(t) - \hat{\phi}(t) \end{bmatrix}$$



State Estimate with Angle Measurement Only

Covariance extrapolation

$$\begin{bmatrix} \dot{p}_{11}(t) & \dot{p}_{12}(t) \\ \dot{p}_{12}(t) & \dot{p}_{22}(t) \end{bmatrix} = \begin{bmatrix} L_p & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} + \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix}\begin{bmatrix} L_p & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} L_p^2\sigma_{p_W}^2 & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{\sigma_{\phi_M}^2}\begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix}^T \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix}$$

Gain computation

$$\begin{bmatrix} k_{11}(t) \\ k_{21}(t) \end{bmatrix} = \frac{1}{\sigma_{\phi_M}^2}\begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix}$$

State estimate with roll angle measurement

$$\begin{bmatrix} \dot{\hat{p}}(t) \\ \dot{\hat{\phi}}(t) \end{bmatrix} = \begin{bmatrix} L_p & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \hat{p}(t) \\ \hat{\phi}(t) \end{bmatrix} + \begin{bmatrix} L_{\delta A} \\ 0 \end{bmatrix}\delta A(t) + \begin{bmatrix} k_{11}(t) \\ k_{21}(t) \end{bmatrix}\left[\phi_M(t) - \hat{\phi}(t)\right]$$

State and Output Vectors for the Quadrotor Helicopter

Longitudinal State

$$\mathbf{x} = \begin{bmatrix} x \\ z \\ u \\ w \\ q \\ \theta \end{bmatrix} \begin{matrix} Range \\ Height \\ Axial\ Velocity \\ Normal\ Velocity \\ Pitch\ Rate \\ Pitch\ Angle \end{matrix}$$

Longitudinal Output

$$\mathbf{y} = \begin{bmatrix} x \\ -z \\ -z \\ -z \\ \dot{x} \\ \dot{u} \\ \dot{w} \\ q \\ \theta \end{bmatrix} \begin{matrix} Range(GPS) \\ Height(GPS) \\ Height(Ultrasound) \\ Height(Pressure\ Sensor) \\ Ground\ Speed(QVGA\ Camera) \\ Axial\ Acceleration \\ Normal\ Acceleration \\ Pitch\ Rate(Gyro) \\ Pitch\ Angle(Magnetometer) \end{matrix}$$

$$\mathbf{y} \neq \mathbf{x}$$

Output Vector and Matrix for the Quadrotor Helicopter

Neglect GPS and Pressure Sensor

Longitudinal Output, Linearized at $\theta = 0$

$$\mathbf{y} = \mathbf{H_x x} + \mathbf{H_u u}$$

$$= \begin{bmatrix} -z \\ \dot{x} \\ \dot{u} \\ \dot{w} \\ q \\ \theta \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -g \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ -2/m & -2/m \\ d/I_{yy} & -d/I_{yy} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} T_f \\ T_a \end{bmatrix}$$

How would you design the Kalman Filter?