Adaptive State Estimation

Robert Stengel
Optimal Control and Estimation MAE 546
Princeton University, 2015

- Nonlinearity of adaptation
- Parameter-adaptive filtering
- Test for whiteness of the residual
- Bias estimation and noise-adaptive filtering
- Multiple model estimation

http://www.princeton.edu/~stengel/MAE546.html
http://www.princeton.edu/~stengel/OptConEst.html

Nonlinearity of Adaptation

- Adaptation required if
  - System parameters are unknown
  - System structure is unknown
  - Disturbance/measurement statistics are uncertain

- Adaptive estimators are fundamentally nonlinear, even if the system is linear
  - Parameters to be estimated multiply the state
  - Statistics are derived from measurement residuals
  - Estimator gain depends on parameter estimates

"Process Noise"

White noise disturbance input ("process noise") is similar to random parameter variation

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{F}(\mathbf{p})\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}\mathbf{w}(t)\\ &\simeq \mathbf{F}(\mathbf{p}_o)\mathbf{x}(t) + \left\{\frac{\partial \mathbf{F}(\mathbf{p}_o)}{\partial \mathbf{p}}\Delta\mathbf{p}(t)\right\}\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \mathbf{L}\mathbf{w}(t)\\ &\triangleq \mathbf{F}(\mathbf{p}_o)\mathbf{x}(t) + \mathbf{G}\mathbf{u}(t) + \begin{bmatrix} \mathbf{L}_{\mathbf{w}} & \mathbf{L}_{\Delta\mathbf{p}}\left[\mathbf{x}(t)\right] \end{bmatrix}\begin{bmatrix} \mathbf{w}(t) \\ \Delta\mathbf{p}(t) \end{bmatrix}\end{aligned}$$

$\left\{\frac{\partial \mathbf{F}(\mathbf{p}_o)}{\partial \mathbf{p}}\Delta\mathbf{p}(t)\right\}$ is an $(n \times n)$ matrix

$\mathbf{L}_{\Delta\mathbf{p}}\left[\mathbf{x}(t)\right]$ is $(n \times s)$ matrix

$$\mathbf{w}'(t) \triangleq \begin{bmatrix} \mathbf{w}(t) \\ \Delta\mathbf{p}(t) \end{bmatrix}; \quad E\left[\mathbf{w}'(t)\right] = \mathbf{0}; \quad E\left\{\mathbf{w}'(t)\mathbf{w}'^T(\tau)\right\} = \mathbf{Q}'\delta(t-\tau)$$

… however, model is approximate and nonlinear

Parameter-Adaptive Estimation (Parameter Identification via Extended Kalman Filter)

Parameter-Dependent Linear System

Linear system with parameter-dependent sensitivity matrices

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

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

$$E\begin{bmatrix} \mathbf{w}(t) \\ \mathbf{n}(t) \end{bmatrix} = \mathbf{0}; \quad E\left\{\begin{bmatrix} \mathbf{w}(t) \\ \mathbf{n}(t) \end{bmatrix}\begin{bmatrix} \mathbf{w}^T(\tau) & \mathbf{n}^T(\tau) \end{bmatrix}\right\} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix}\delta(t-\tau)$$

Parameter vector, $\mathbf{p}(t)$, could be

Known: a prescribed function of time

Unknown: the output of a random dynamic process

Dynamic Model for Parameter Estimation

- Augment vector to include original state and parameter vector

$$\boxed{\mathbf{x}_A(t) \triangleq \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{p}(t) \end{bmatrix}}$$

- Augment system model for parameter identification
- System is nonlinear because parameter is contained in the augmented state

$$\dot{\mathbf{x}}_A(t) = \left\{\begin{array}{c} \left[\mathbf{F}[\mathbf{p}(t)]\mathbf{x}(t) + \mathbf{G}[\mathbf{p}(t)]\mathbf{u}(t) + \mathbf{L}[\mathbf{p}(t)]\mathbf{w}_x(t)\right] \\ \mathbf{f}_p[\mathbf{p}(t), \mathbf{w}_p(t)] \end{array}\right\} \triangleq \mathbf{f}_A[\mathbf{x}(t), \mathbf{p}(t), \mathbf{u}(t), \mathbf{w}(t)]$$

$$\mathbf{z}(t) = \mathbf{H}_A[\mathbf{p}(t)]\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{p}(t) \end{bmatrix} + \mathbf{n}(t) = \left\{ \mathbf{H}[\mathbf{p}(t)] \quad \mathbf{0} \right\}\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{p}(t) \end{bmatrix} + \mathbf{n}(t) + \mathbf{b}[\mathbf{p}(t)]$$

$\mathbf{H}[\mathbf{p}(t)]$: Unknown scale factors and coupling terms (TBD)

$\mathbf{b}[\mathbf{p}(t)]$: Unknown bias error (TBD)

Parameter Vector Must Have a Dynamic Model

Unknown constant: p(*t*) = constant

$$\boxed{\dot{\mathbf{p}}(t)=\mathbf{0};\quad \mathbf{p}(0)=\mathbf{p}_o;\quad \mathbf{P}_{\mathbf{p}}(0)=\mathbf{P}_{\mathbf{p}_o}}$$

Random p(*t*) (integrated white noise)

$$\begin{aligned}&\dot{\mathbf{p}}(t)=\mathbf{w}_{\mathbf{p}}(t);\quad \mathbf{p}(0)=\mathbf{0};\quad \mathbf{P}_{\mathbf{p}}(0)=\mathbf{P}_{\mathbf{p}_o}\\ &E\left[\mathbf{w}_{\mathbf{p}}(t)\right]=\mathbf{0};\quad E\left[\mathbf{w}_{\mathbf{p}}(t)\mathbf{w}_{\mathbf{p}}^T(\tau)\right]=\mathbf{Q}_{\mathbf{p}}\delta(t-\tau)\end{aligned}$$

Linear dynamic system (Markov process)

$$\boxed{\dot{\mathbf{p}}(t)=\mathbf{A}\mathbf{p}(t)+\mathbf{B}\mathbf{w}_{\mathbf{p}}(t)\triangleq\mathbf{f}_{\mathbf{p}}\left[\mathbf{p}(t),\mathbf{w}_{\mathbf{p}}(t)\right];\quad \mathbf{w}_{\mathbf{p}}(t)\sim N\left(\mathbf{0},\mathbf{Q}_{\mathbf{p}}\right)}$$

Dynamic Models for the Parameter Vector

Doubly integrated white noise

$$\boxed{\dot{\mathbf{p}}_M(t)=\begin{bmatrix}\dot{\mathbf{p}}(t)\\ \dot{\mathbf{p}}_D(t)\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}(t)\\ \mathbf{p}_D(t)\end{bmatrix}+\begin{bmatrix}\mathbf{0}\\ \mathbf{w}_{\mathbf{p}}(t)\end{bmatrix}}$$

$$\begin{bmatrix}\mathbf{p}(t)\\ \mathbf{p}_D(t)\end{bmatrix}=\begin{bmatrix}\text{Parameter vector}\\ \text{Parameter rate of change}\end{bmatrix}$$

Triply integrated white noise

$$\boxed{\dot{\mathbf{p}}_M(t)=\begin{bmatrix}\dot{\mathbf{p}}(t)\\ \dot{\mathbf{p}}_D(t)\\ \dot{\mathbf{p}}_A(t)\end{bmatrix}=\begin{bmatrix}\mathbf{0} & \mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}\\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}(t)\\ \mathbf{p}_D(t)\\ \mathbf{p}_A(t)\end{bmatrix}+\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{w}_{\mathbf{p}}(t)\end{bmatrix}}$$

$$\begin{bmatrix}\mathbf{p}(t)\\ \mathbf{p}_D(t)\\ \mathbf{p}_A(t)\end{bmatrix}=\begin{bmatrix}\text{Parameter vector}\\ \text{Parameter rate of change}\\ \text{Parameter acceleration}\end{bmatrix}$$

Weathervane Example (4.7-1)

2^{nd} - order system

Constant parameter, $\omega_n^2 = a \equiv 4$,

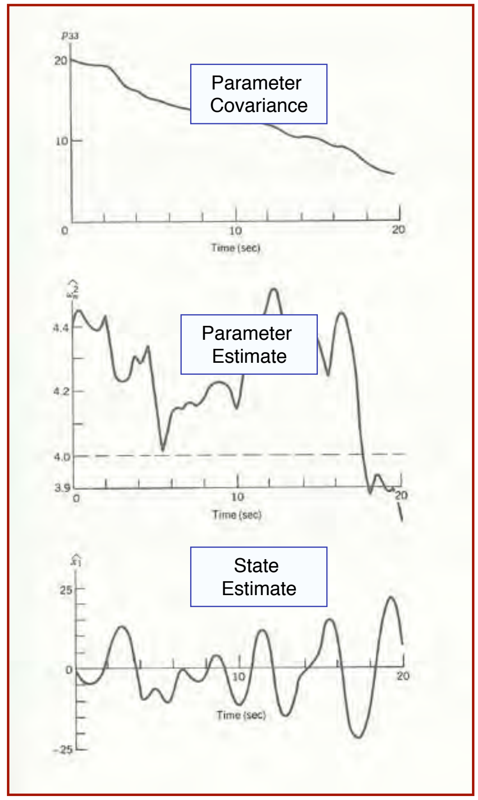
Assumed to be 4.4

$b = 0.4$ **(known)**

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{a} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ a & b & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ a \end{bmatrix} + \begin{bmatrix} 0 \\ -a \\ 0 \end{bmatrix} w$$

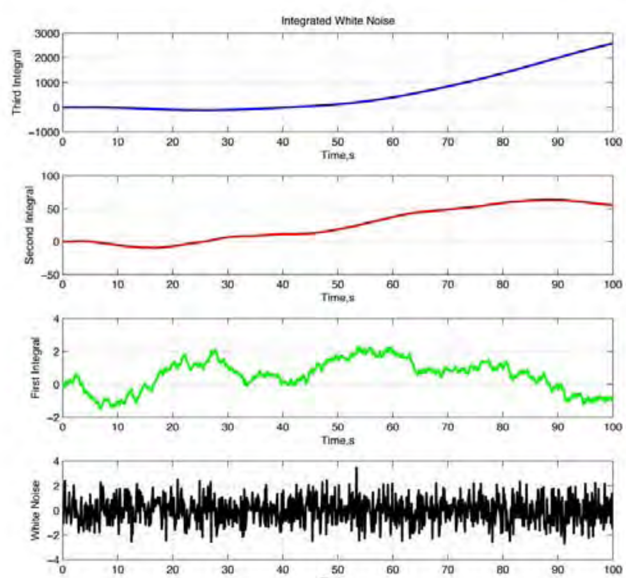
$$Q = 1000; \quad \mathbf{R} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}; \quad \mathbf{P}_p(0) = 20$$

Additional details in text



Integrated White Noise Models of a Parameter

- **Third integral models slowly varying, smooth parameter**
- **Second integral is smoother but still has fast changes**
- **First integral of white noise has abrupt jumps, valleys, and peaks**
- **White noise**



Hybrid Filter for Parameter Estimation

Extrapolation of Augmented State

$$\hat{\mathbf{x}}_A\left[t_k(-)\right] = \mathbf{x}_A\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \mathbf{f}_A\left[\hat{\mathbf{x}}_A(\tau), \mathbf{u}(\tau)\right] d\tau$$

Covariance Extrapolation

$$\mathbf{P}_A\left[t_k(-)\right] = \mathbf{P}_A\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \left[\mathbf{F}_A(\tau)\mathbf{P}_A(\tau) + \mathbf{P}_A(\tau)\mathbf{F}_A^T(\tau) + \mathbf{L}_A(\tau)\mathbf{Q}'_C(\tau)\mathbf{L}_A^T(\tau)\right] d\tau$$

Filter Gain Calculation

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

$(\mathbf{F}_A, \mathbf{H}_A)$ must be locally observable except at isolated points

Hybrid Filter for Parameter Estimation

State Update

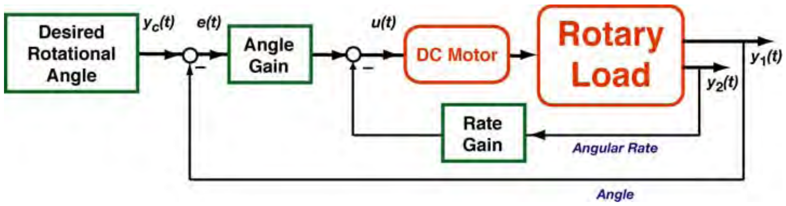
$$\hat{\mathbf{x}}_A\left[t_k(+)\right] = \hat{\mathbf{x}}_A\left[t_k(-)\right] + \mathbf{K}(t_k)\left\langle \mathbf{z}(t_k) - \mathbf{h}\left\{\hat{\mathbf{x}}\left[t_k(-)\right]\right\}\right\rangle$$

Covariance Update

$$\mathbf{P}_A\left[t_k(+)\right] = \left[\mathbf{I}_n - \mathbf{K}(t_k)\mathbf{H}_A(t_k)\right]\mathbf{P}_A\left[t_k(-)\right]$$

$(\mathbf{F}_A, \mathbf{H}_A)$ must be locally observable except at isolated points

Example: Estimation of Variable Rotary Load for a Robot Arm Elbow



Closed-loop dynamic equation

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -c_1 / J(t) & -c_2 / J(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ c_1 / J(t) \end{bmatrix} y_c$$

Parameter, $p(t) \triangleq$ Rotary load inertia, $J(t)$; c_1, c_2 : Control gains (given)

Augmented State and Measurement for Unknown Inertia Modeled as Doubly Integrated Parameter

$$\mathbf{x}_A(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ J(t) \\ \dot{J}(t) \end{bmatrix} \triangleq \begin{bmatrix} x_1(t) \\ x_2(t) \\ p(t) \\ p_D(t) \end{bmatrix} \quad \begin{matrix} \text{Angle} \\ \text{Angular Rate} \\ \text{Rotary Load Inertia} \\ \text{Inertia Rate} \end{matrix}$$

$$\mathbf{z}(t) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ p(t) \\ p_D(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \end{bmatrix}$$

Nonlinear Dynamic Equation with Unknown Inertia Modeled as Doubly Integrated White Noise

$$\dot{\mathbf{x}}_A(t) = \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{p}(t) \\ \dot{p}_D(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ -c_1/p(t) & -c_2/p(t) & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ p(t) \\ p_D(t) \end{bmatrix} + \begin{bmatrix} 0 \\ c_1/p(t) \\ 0 \\ 0 \end{bmatrix} y_c + \begin{bmatrix} 0 \\ w_2(t) \\ 0 \\ w_{p_D}(t) \end{bmatrix} \triangleq \mathbf{f}_A[\bullet]$$

Stability and Measurement Matrices

$$\mathbf{F}_A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ -c_1/\hat{p}(t) & -c_2/\hat{p}(t) & \left[\dfrac{c_1}{2\hat{p}^2(t)}\hat{x}_1(t) + \dfrac{c_2}{2\hat{p}^2(t)}\hat{x}_2(t)\right] & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{H}_A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Means and Covariances

$$\hat{\mathbf{x}}_A(0) = E\begin{bmatrix} x_1(0) \\ 0 \\ p(0) \\ 0 \end{bmatrix} \triangleq \begin{bmatrix} \hat{x}_1(0) \\ 0 \\ \hat{p}(0) \\ 0 \end{bmatrix}$$

$$\mathbf{P}_A(0) = E\left\{\left[\mathbf{x}_A(0) - \hat{\mathbf{x}}_A(0)\right]\left[\mathbf{x}_A(0) - \hat{\mathbf{x}}_A(0)\right]^T\right\}$$

$$\hat{\mathbf{w}}(t) = \mathbf{0}; \quad \mathbf{Q}(t) = E\left[\mathbf{w}(t)\mathbf{w}^T(\tau)\right]$$

$$\hat{\mathbf{n}}(t) = \mathbf{0}; \quad \mathbf{R}(t) = E\left[\mathbf{n}(t)\mathbf{n}^T(\tau)\right]$$

Hybrid Filter for Robot Elbow State and Load Estimation

Convergence is problem-dependent

Qualitative observability of parameter

Actual and assumed uncertainty covariances

Accuracy of dynamic model

Extrapolation of Augmented State

$$\hat{\mathbf{x}}_A\left[t_k(-)\right] = \mathbf{x}_A\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \mathbf{f}_A\left[\hat{\mathbf{x}}_A(\tau), \mathbf{u}(\tau)\right] d\tau$$

Covariance Extrapolation

$$\mathbf{P}_A\left[t_k(-)\right] = \mathbf{P}_A\left[t_{k-1}(+)\right] + \int_{t_{k-1}}^{t_k} \left[\mathbf{F}_A(\tau)\mathbf{P}_A(\tau) + \mathbf{P}_A(\tau)\mathbf{F}_A^T(\tau) + \mathbf{L}_A(\tau)\mathbf{Q}'_C(\tau)\mathbf{L}_A^T(\tau)\right] d\tau$$

Filter Gain Calculation

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

State Update

$$\hat{\mathbf{x}}_A\left[t_k(+)\right] = \hat{\mathbf{x}}_A\left[t_k(-)\right] + \mathbf{K}(t_k)\left\langle \mathbf{z}(t_k) - \mathbf{h}\left\{\hat{\mathbf{x}}\left[t_k(-)\right]\right\}\right\rangle$$

Covariance Update

$$\mathbf{P}_A\left[t_k(+)\right] = \left[\mathbf{I}_n - \mathbf{K}(t_k)\mathbf{H}_A(t_k)\right]\mathbf{P}_A\left[t_k(-)\right]$$

Bias Estimation and Noise-Adaptive Filtering

Residuals and Optimal Filtering

Linear-optimal filtering has the innovations property

Optimal estimation extracts all the available information from the measurements

Measurement residual should be zero-mean white noise

State estimate should be orthogonal to the error

Residual and its statistics

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

$$E\left[\mathbf{r}_k(-)\right] = \hat{\mathbf{r}}_k(-) = \mathbf{0}$$

$$E\left[\mathbf{r}_k(-)\mathbf{r}_k^T(-)\right] \triangleq \mathbf{S}_k(-) = \mathbf{HP}_k(-)\mathbf{H}^T + \mathbf{R}_k$$

$$E\left(\hat{\mathbf{x}}_k\mathbf{n}_k^T\right) = \mathbf{0}$$

Residual Should Be White Noise if State Estimate is Optimal

Test for whiteness using <u>normalized autocovariance function</u>

Sampled (batch process) estimate of the autocovariance function matrix, C(k)

$$\mathbf{C}(k) = \left(\frac{1}{N}\right)\sum_{n=k}^{N} \mathbf{r}_n \mathbf{r}_{n+k}^T, \quad k << N$$

$$\dim\left[\mathbf{C}(k)\right] = r \times r$$

Normalize diagonal elements of C(k) by their zero-lag (k = 0) values

$$\rho_{ij}(k) = \frac{c_{ij}(k)}{c_{ii}(0)}$$

Test for Residual Whiteness

If $\mathbf{r}_k$ is white

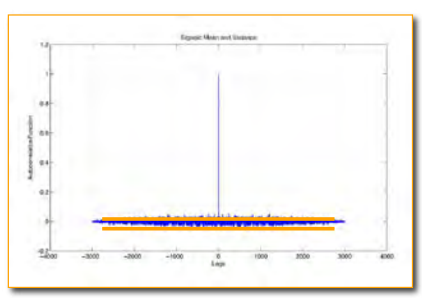
$$\rho_{ij}(k) = \begin{cases} 1, & i = j \text{ and } k = 0 \\ 0, & i \neq j \text{ or } k \neq 0 \end{cases}, \quad N \to \infty$$

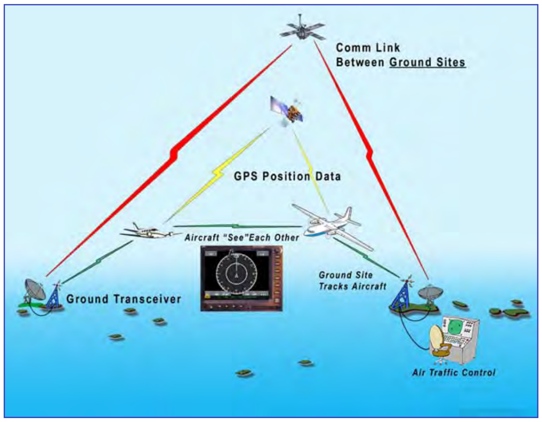
Off-diagonal terms should be negligible

For finite sample, 95% confidence limit based on diagonal elements

$$\left|\rho_{ii}(k)\right| \leq \frac{1.96}{\sqrt{N}}, \quad k \neq 0$$

Test is passed if 19 out of 20 non-zero-lag samples are within the limit





Bias Example: Advanced Dependent Surveillance (ADS-B) System for Air Traffic Control

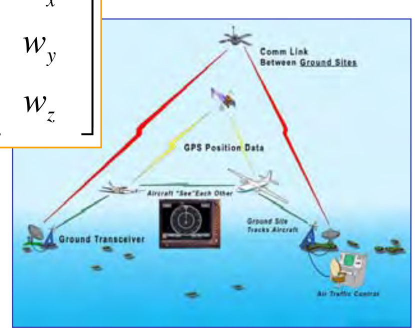
- Surveillance and tracking radars on ground
- GPS/Inertial/Air data measurements in aircraft
- Satellite/Line-of-sight communications links
- Air traffic control centers (ATCC)
  - Prevent collisions
  - Maintain efficient flow of air traffic

Dynamics of Individual Aircraft

Surveillance equations of motion

Neglect fast dynamics of aircraft

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{v}_x \\ \dot{v}_y \\ \dot{v}_z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ v_x \\ v_y \\ v_z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ w_x \\ w_y \\ w_z \end{bmatrix}$$



ATCC Measurements and Communicated Locations of Individual Aircraft

$$\mathbf{z} = \begin{bmatrix} \text{Ground-based radar measurements} \\ \text{Aircraft} \begin{cases} \text{GPS measurements} \\ \text{Inertial navigation measurements} \\ \text{Air data measurements} \end{cases} \end{bmatrix} = (\mathbf{Hx} + \mathbf{n}) + \mathbf{b}$$

$$\mathbf{b} = \color{red}{\textbf{Biases}} = \begin{bmatrix} \text{Radar Range Bias} \\ \text{Radar Azimuth Bias} \\ \text{Aircraft Altimeter Bias} \\ \text{ADS-B Longitude Bias} \\ \text{ADS-B Latitude Bias} \\ \text{ADS-B Altitude Bias} \\ \text{ADS-B North Velocity Bias} \\ \text{ADS-B East Velocity Bias} \end{bmatrix}$$

Bias is a quasi-constant error that is not related to random noise
Least-squares estimator does not reduce bias error

Measurement Bias and Covariance Estimation

Batch processing estimates

$$\bar{\mathbf{r}}(-) \triangleq \hat{\mathbf{b}}(-) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{r}_i(-) = \frac{1}{N}\sum_{i=1}^{N}\left[\mathbf{z}_i - \mathbf{H}\hat{\mathbf{x}}_i(-)\right]$$ **Bias Estimate**

$$\hat{\mathbf{S}} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{r}_i - \bar{\mathbf{r}}_i)(\mathbf{r}_i - \bar{\mathbf{r}}_i)^T$$ **Sample Covariance Matrix**

$$\hat{\mathbf{R}} = \hat{\mathbf{S}} - \frac{N-1}{N}\mathbf{H}\mathbf{P}_k(-)\mathbf{H}^T$$ **Measurement Noise Estimate**

Running Estimate of Measurement Bias and Error Covariance

$$\begin{aligned}\hat{\mathbf{b}}_k &= \hat{\mathbf{b}}_{k-1} + k_{bias}\left\{\left[\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k(-)\right] - \hat{\mathbf{b}}_{k-1}\right\} \\ &= \hat{\mathbf{b}}_{k-1} + k_{bias}\left[\mathbf{r}_k - \hat{\mathbf{b}}_{k-1}\right] \\ &= \left[1 - k_{bias}\right]\hat{\mathbf{b}}_{k-1} + k_{bias}\mathbf{r}_k\end{aligned}$$

$$k_{bias}, k_{noise} < 1$$

$$\hat{\mathbf{R}}_k = \hat{\mathbf{R}}_{k-1} + k_{noise}\left\{\left[\left(\mathbf{r}_k - \hat{\mathbf{b}}_k\right)\left(\mathbf{r}_k - \hat{\mathbf{b}}_k\right)^T - \mathbf{H}\mathbf{P}_k(-)\mathbf{H}^T\right] - \hat{\mathbf{R}}_{k-1}\right\}$$

Options

- Choose add hoc recursive gain
- Use weighted least-squares estimator
- Incorporate in an integrated parameter-adaptive filter

Disturbance Bias Estimation

System equation

$$\mathbf{x}_{k+1} = \mathbf{\Phi}\mathbf{x}_k + \mathbf{w}_k$$

Disturbance residual

$$\mathbf{w}_k = \mathbf{\Phi}\mathbf{x}_k - \mathbf{x}_{k+1} \approx \mathbf{\Phi}\hat{\mathbf{x}}_k(+) - \hat{\mathbf{x}}_{k+1}(+)$$

Sample mean

$$\bar{\mathbf{w}} \triangleq \hat{\mathbf{w}}(+) = \frac{1}{N}\sum_{i=1}^{N}\left[\mathbf{\Phi}\hat{\mathbf{x}}_i(+) - \hat{\mathbf{x}}_{i+1}(+)\right] = \frac{1}{N}\sum_{i=1}^{N}\mathbf{w}_i$$

Disturbance Covariance Estimation

Sample covariance

$$\hat{\mathbf{W}} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\mathbf{w}_i - \bar{\mathbf{w}}\right)\left(\mathbf{w}_i - \bar{\mathbf{w}}\right)^T$$

Disturbance covariance estimate

$$\hat{\mathbf{Q}} = \hat{\mathbf{W}} - \frac{N-1}{N}\mathbf{\Phi}\mathbf{P}_k(+)\mathbf{\Phi}^T$$

Running Estimate of Disturbance Bias and Covariance

$$\hat{\mathbf{w}}_k(+) = \hat{\mathbf{w}}_{k-1}(+) + k_{bias}\left\{\mathbf{w}_k(+) - \hat{\mathbf{w}}_{k-1}(+)\right\}$$

$$\hat{\mathbf{Q}}_k = \hat{\mathbf{Q}}_{k-1} + k_{noise}\left\{\left[\left(\mathbf{w}_k - \hat{\mathbf{w}}_k\right)\left(\mathbf{w}_k - \hat{\mathbf{w}}_k\right)^T - \frac{N-1}{N}\mathbf{\Phi}\mathbf{P}_k(+)\mathbf{\Phi}^T\right] - \hat{\mathbf{Q}}_{k-1}\right\}$$

$$k_{bias}, k_{noise} < 1$$

- Options as before
  - Choose add hoc recursive gain
  - Use weighted least-squares estimator
  - Incorporate in an integrated parameter-adaptive filter

Noise-and-Bias Adaptive Filter

Use separately estimated means and covariances in Kalman filter

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

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

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

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

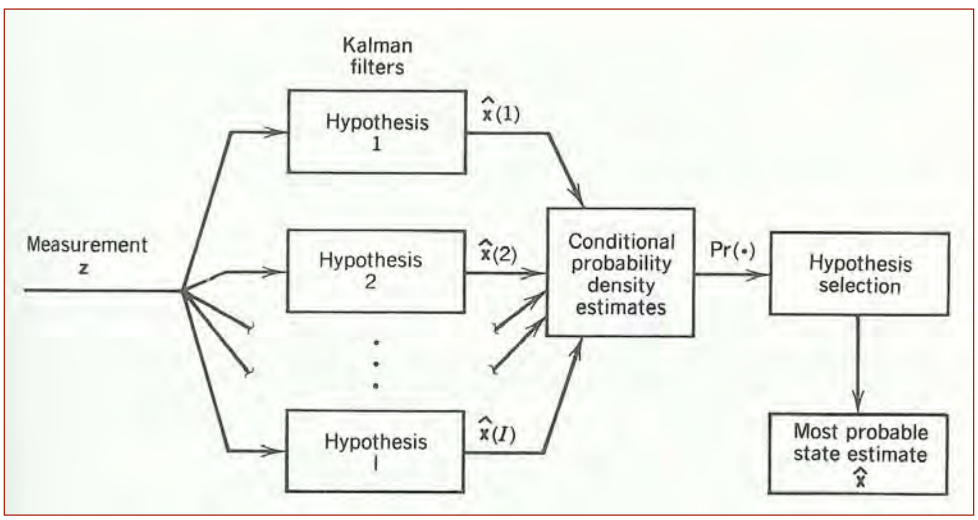
$$\boxed{\mathbf{P}_k(+) = \left[\mathbf{P}_k^{-1}(-) + \mathbf{H}_k^T\hat{\mathbf{R}}_k^{-1}\mathbf{H}_k\right]^{-1}}$$

Multiple Model Estimation

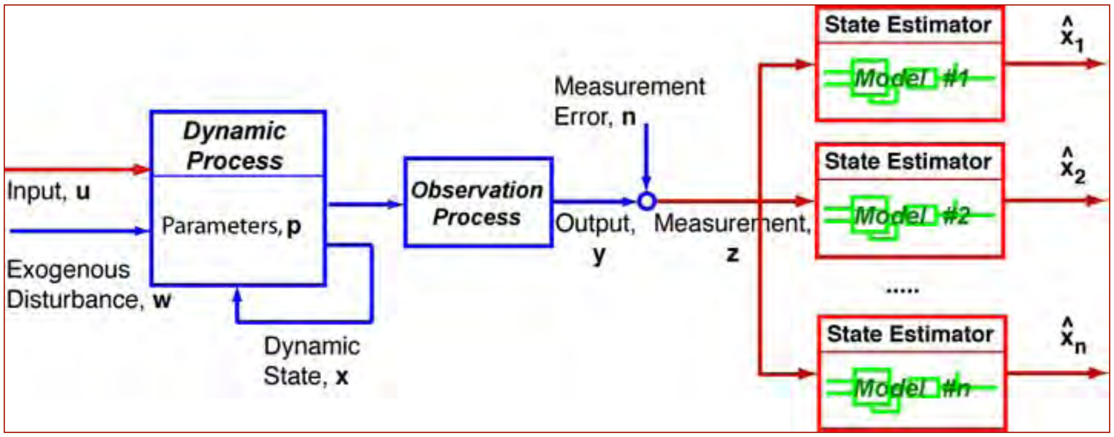
Multiple Model Estimation

- Bank of Kalman filters, each "tuned" to a different hypothesis
  - Different model parameters or structures
  - Different uncertainty models
  - Different initial conditions
- Best performance determined by a hypothesis test, e.g., Maximum Likelihood
- State estimate chosen accordingly

Hypothesis Testing



Multiple Model Estimation



Multiple Model Estimation

- Consider ***J* systems** distinguished by ***J*** parameter vectors
- Conditional probability mass function for the ***j^th^*** parameter set

$$\boxed{\Pr\left(\mathbf{p}_j \mid \mathbf{z}_k\right) = \frac{\mathrm{pr}\left(\mathbf{z}_k \mid \mathbf{p}_j\right)\Pr\left(\mathbf{p}_j\right)_{k-1}}{\sum_{i=1}^{I}\left[\mathrm{pr}\left(\mathbf{z}_k \mid \mathbf{p}_i\right)\Pr\left(\mathbf{p}_i\right)_{k-1}\right]}}$$

- Probability that the measurement at ***k* – 1** was obtained is one; therefore,

$$\boxed{\Pr\left(\mathbf{p}_j\right)_{k-1} = \Pr\left(\mathbf{p}_j \mid \mathbf{z}_{k-1}\right)}$$

- ... and the equation forms the basis for a recursion

$$\boxed{\Pr\left(\mathbf{p}_j \mid \mathbf{z}_k\right) = \frac{\mathrm{pr}\left(\mathbf{z}_k \mid \mathbf{p}_j\right)}{\sum_{i=1}^{J}\left[\mathrm{pr}\left(\mathbf{z}_k \mid \mathbf{p}_i\right)\Pr\left(\mathbf{p}_i\right)_{k-1}\right]}\Pr\left(\mathbf{p}_j \mid \mathbf{z}_{k-1}\right)}$$

Multiple Model Estimation

Conditional probability density function $\mathbf{z}_k$ must be found

With the true parameter set

$$\boxed{\begin{gathered}\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{n}_k \\ \mathbf{x}_k = \mathbf{\Phi}\mathbf{x}_{k-1} + \mathbf{\Gamma}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}, \quad \mathbf{x}_o = \mathbf{x}(0)\end{gathered}}$$

If the true state were known

$$\boxed{\begin{aligned}\mathrm{pr}(\mathbf{z}_k \mid \mathbf{p}) &= \mathrm{pr}\left[\mathbf{z}_k \mid \mathbf{x}_k(\mathbf{p})\right] \\ &= \frac{1}{(2\pi)^{n/2}\left|\mathbf{R}_k\right|^{1/2}} e^{-\frac{1}{2}(\mathbf{z}_k - \mathbf{H}\mathbf{x}_k)^T \mathbf{R}_k^{-1}(\mathbf{z}_k - \mathbf{H}\mathbf{x}_k)} \\ &= \frac{1}{(2\pi)^{n/2}\left|\mathbf{R}_k\right|^{1/2}} e^{-\frac{1}{2}\mathbf{n}_k{}^T \mathbf{R}_k^{-1}\mathbf{n}_k}\end{aligned}}$$

However, only an estimate of $\mathbf{x}_k$ is available

Multiple Model Estimation

- **The state is estimated by a Kalman filter for the "true" parameter**

$$\boxed{\mathrm{pr}\left[\mathbf{z}_k \mid \hat{\mathbf{x}}_k(\mathbf{p})\right] = \frac{1}{(2\pi)^{n/2}\left|\mathbf{S}_k\right|^{1/2}} e^{-\frac{1}{2}\mathbf{r}_k^T \mathbf{S}_k^{-1}\mathbf{r}_k}}$$

with

$$\boxed{\begin{aligned}\mathbf{r}_k(-) &= \mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k(-) \\ \mathbf{S}_k &= \mathbf{H}\mathbf{P}_k(-)\mathbf{H}^T + \mathbf{R}_k\end{aligned}}$$

- **The bank of *J* Kalman filters is formed with each filter assuming that different parameters are the <u>true</u> parameter**

Conditional Probabilities and the Adaptive State Estimate

Conditional probabilities for each hypothesis

$$\Pr\left(\mathbf{p}_j \mid \mathbf{z}_k\right) = \frac{\mathrm{pr}\left[\mathbf{z}_k \mid \hat{\mathbf{x}}_k\left(\mathbf{p}_j\right)\right]\Pr\left(\mathbf{p}_j \mid \mathbf{z}_{k-1}\right)}{\sum_{i=1}^{J}\left\{\mathrm{pr}\left[\mathbf{z}_k \mid \hat{\mathbf{x}}_k\left(\mathbf{p}_i\right)\right]\Pr\left(\mathbf{p}_i \mid \mathbf{z}_{k-1}\right)\right\}}, \quad j = 1, J$$

State estimate is chosen to be

the one for which the conditional probability is highest, or a weighted sum of the state estimates

$$\hat{\mathbf{x}}_k(+) = \sum_{i=1}^{J}\left\{\Pr\left(\mathbf{p}_i \mid \mathbf{z}_k\right)\hat{\mathbf{x}}_k\left[\mathbf{p}_i,(+)\right]\right\}$$

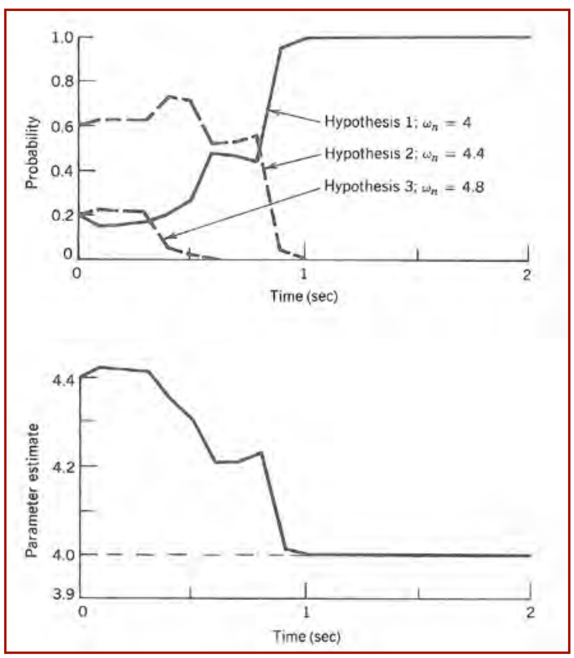
Parameter vector is chosen accordingly

Weathervane Example (4.7-2)

- **2^{nd}-order system with three hypothesized natural frequencies**

$$\omega_n^2 = \begin{cases} 4 & [Correct] \\ 4.4 & [Expected] \\ 4.8 & \end{cases}$$
$$\zeta = 0.1$$

- **Algorithm searches, then homes in on correct solution**

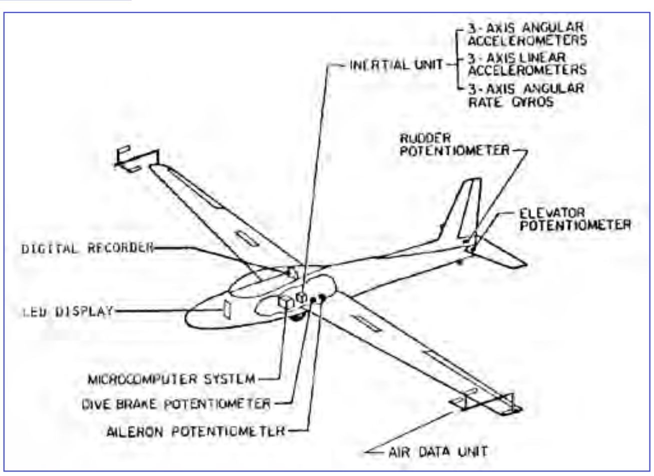


Next Time: Stochastic Optimal Control

Supplemental Material

Estimating Parameters of Nonlinear Systems

Aerodynamic Coefficients of a Sailplane from Flight Data*



- Princeton University Flight Research Laboratory

*** Sri-Jayantha and Stengel, 1988**

Dynamic Equations of the Sailplane

Body-axis velocity and angular rate equations using quaternions

$$
\begin{bmatrix} \dot{x}_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} \dot{u} \\ v \\ w \\ p \\ q \\ r \end{bmatrix} = \begin{bmatrix} rv - qw + 2g(e_2e_4 - e_1e_3) + X \\ pw - ru + 2g(e_2e_3 + e_1e_4) + Y \\ qu - pv + g(e_1^2 + e_2^2 - e_3^2 - e_4^2) + Z \\ pqC_1 + qrC_2 + qC_3 + L + NC_4 \\ prC_5 + (r^2 - p^2)C_6 - rC_7 + M \\ pqC_8 + qrC_9 + qC_{10} + LC_{11} + N \end{bmatrix}
$$

Propagation of quaternions from angular rates

$$
\begin{pmatrix} \dot{x}_7 \\ x_8 \\ x_9 \\ x_{10} \end{pmatrix} = \begin{pmatrix} \dot{e}_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix} = (1/2) \begin{pmatrix} 0 & -r & -q & -p \\ r & 0 & -p & q \\ q & p & 0 & -r \\ p & -q & r & 0 \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}
$$

Definitions of Terms

Definitions of terms in dynamic equations

$X = qSC_X/m,\ Y = qSC_Y/m,\ Z = qSC_Z/m$

$L = qSb(C_l/I_{XX})\{I_{XX}I_{ZZ}/(I_{XX}I_{ZZ} - I_{XZ}^2)\}$

$M = qSc(C_m/I_{YY}) + \{(Xm)_z - Zm)_x\}/I_{YY}\}$

$N = qSb(C_n/I_{ZZ})\{I_{XX}I_{ZZ}/(I_{XX}I_{ZZ} - I_{XZ}^2)\}$

$C_1 = \{I_{XZ}(I_{ZZ} + I_{XX} - I_{YY})\}/\Gamma^2$

$C_2 = \{I_{ZZ}(I_{YY} - I_{ZZ}) - I_{XZ}^2\}/\Gamma^2$

$C_3 = 0$ (case with no rotating engine components)

$C_4 = I_{XZ}/I_{XX}$

$C_5 = (I_{ZZ} - I_{XX})/I_{YY}$

$C_6 = I_{XZ}/I_{YY}$

$C_7 = 0$

$C_8 = \{I_{XX}(I_{XX} - I_{YY}) + I_{XZ}^2\}/\Gamma^2$

$C_9 = I_{XZ}(I_{YY} - I_{ZZ} - I_{XX})/\Gamma^2$

$C_{10} = 0$

$C_{11} = I_{XZ}/I_{ZZ}$

$q = (1/2)\rho V^2,\ \Gamma^2 = (I_{XX}I_{ZZ} - I_{XZ}^2)$

Specific force and moment definitions for parameter identification

$x_{12} = b_{10} = X =$ Axial specific force, ft/s^2

$x_{15} = b_{20} = Y =$ Side specific force, ft/s^2

$x_{18} = b_{30} = Z =$ Normal specific force, ft/s^2

$x_{21} = b_{40} = L =$ Roll specific moment, rad/s^2

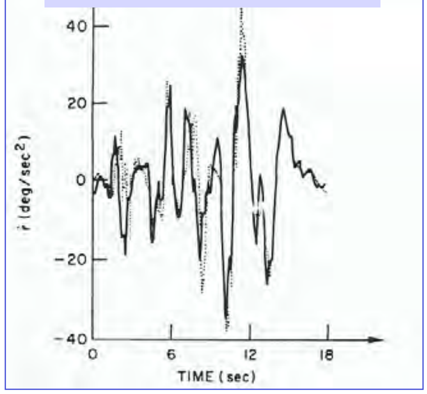
$x_{24} = b_{50} = M =$ Pitch specific moment, rad/s^2

$x_{27} = b_{60} = N =$ Yaw specific moment, rad/s^2

Aerodynamic Coefficients of a Sailplane from Flight Data

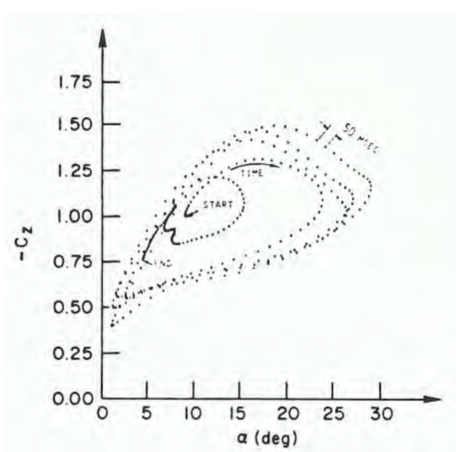
- Estimation-before-modeling technique
  - Estimate the state
    - Extended Kalman Filter (forward pass)
    - Modified Bryson-Frazier Smoother (backward pass)
  - Use multivariate regression to model aerodynamic coefficients

Measured vs. Estimated Yaw Rate



Aerodynamic Coefficients of a Sailplane from Flight Data

Smoothed Estimate of Normal-Force Coefficient (C_Z) History



Estimate of Lift Coefficient (C_L) vs. Angle of Attack