Linear-Optimal Estimation for Continuous-Time Systems
Robert Stengel
Optimal Control and Estimation MAE 546
Princeton University, 2017

- Propagation of uncertainty in continuous-time systems
- Linear-optimal Gaussian estimator (Kalman-Bucy filter)
- Linear-optimal prediction
- Asymptotic stability of the state estimate
- Duality between estimation and control
- Filter divergence
- Square root filtering
- Correlated disturbance and measurement noise

Copyright 2017 by Robert Stengel. All rights reserved. For educational use only.
http://www.princeton.edu/~stengel/MAE546.html
http://www.princeton.edu/~stengel/OptConEst.html

Uncertain Linear, Time-Varying (LTV) Dynamic Model

- Continuous-time LTV model with known coefficients

$$\dot{x}(t) = F(t)x(t) + G(t)u(t) + L(t)w(t), \quad x(t_o) \text{ given}$$

$$x(t) = x(t_o) + \int_{t_o}^{t} [F(\tau)x(\tau) + G(\tau)u(\tau) + L(\tau)w(\tau)] d\tau$$

$$y(t) = Hx(t) + H_u u(t)$$

- Initial condition and disturbance inputs are not known precisely
- Measurement of state is transformed and is subject to error
Continuous-Time State Estimation

Same general structure as discrete-time estimator

Propagation of Mean Value Estimate in Continuous-Time Systems

Assumed mean value statistics

\[ \begin{align*}
E[x(0)] &= m(0); & E[x(t)] &= m(t); & E[\dot{x}(t)] &= \dot{m}(t) \\
E[w(t)] &= 0; & E[u(t)] &= u(t)
\end{align*} \]

By substitution in the system equation, the expected mean value is

\[ \begin{align*}
E[\dot{x}(t)] &= E[F(t)x(t) + G(t)u(t) + L(t)w(t)] \\
&= F(t)E[x(t)] + G(t)u(t) \\
\dot{m}(t) &= F(t)m(t) + G(t)u(t)
\end{align*} \]
Alternative Derivation of Mean Value Estimate Propagation

The sampled-data case

\[ m(t_k) = \Phi(t_k - t_{k-1}) m(t_{k-1}) + \int_{t_{k-1}}^{t_k} \Phi(t_k - \tau) \left[ G(\tau) u(\tau) \right] d\tau \]

\[ \Phi(\Delta t) = I_n + F \Delta t + \frac{1}{2!} F^2 \Delta t^2 + \frac{1}{3!} F^3 \Delta t^3 + \ldots \]

For small \( \Delta t \)

\[ m(t_k) = \left( I_n + F_{k-1} \Delta t \right) m(t_{k-1}) + \int_{t_{k-1}}^{t_k} \left[ I_n + F_{k-1} (t_k - \tau) \right] \left[ G(\tau) u(\tau) \right] d\tau \]

Mean Value Estimate Propagation

- Rearranging terms

\[ \frac{m(t_k) - m(t_{k-1})}{\Delta t} = F_{k-1} m(t_{k-1}) + \left[ I_n + F_{k-1} \frac{\Delta t}{2} \right] G_{k-1} u_{k-1} \]

- In the limit, the difference equation converges to a differential equation

- Propagation of the expected state

\[ \lim_{\Delta t \to 0} \frac{m(t_k) - m(t_{k-1})}{\Delta t} = \frac{dm(t)}{dt} = \dot{m}(t) = F(t) m(t) + G(t) u(t) \]
Covariance Estimate Propagation

Assumed covariance statistics

\[
E \left\{ \left[ x(t) - m(t) \right] \left[ x(\tau) - m(\tau) \right]^T \right\} = P(t)
\]
\[
E \left[ w(t) w^T(\tau) \right] = Q_c \delta(t - \tau);
E \left\{ \left[ u(t) - \bar{u}(t) \right] \left[ u(\tau) - \bar{u}(\tau) \right]^T \right\} = 0
\]

For small \( \Delta t \)

\[
P_k = (I_n + F_{k-1} \Delta t) P_{k-1} (I_n + F_{k-1} \Delta t)^T + Q_{k-1}
\]
\[
= P_{k-1} + F_{k-1} P_{k-1} \Delta t + (F_{k-1} P_{k-1})^T \Delta t + F_{k-1} P_{k-1} F_{k-1}^T \Delta t^2 + Q_{k-1}
\]

Rearranging

\[
\frac{P_k - P_{k-1}}{\Delta t} = F_{k-1} P_{k-1} + (F_{k-1} P_{k-1})^T + F_{k-1} P_{k-1} F_{k-1}^T \Delta t + \frac{Q_{k-1}}{\Delta t}
\]

Covariance rate of change

As \( \Delta t \rightarrow 0 \), \( t_{k-1} \rightarrow t_k \rightarrow t \)

\[
\lim_{\Delta t \rightarrow 0} \frac{P_k - P_{k-1}}{\Delta t} = \frac{dP(t)}{dt} = \dot{P}(t)
\]

Disturbance uncertainty

\[
Q_{k-1} \approx L(t) Q'_c L^T(t) \Delta t
\]
\[
\therefore \frac{Q_{k-1}}{\Delta t} \rightarrow L(t) Q'_c (t) L^T(t)
\]

\[
\dot{P}(t) = F(t) P(t) + P(t) F(t)^T + L(t) Q'_c (t) L^T(t)
\]
Kalman-Bucy Filter

- Optimal estimator for linear systems with Gaussian uncertainty
- Three equations
  1) State estimate extrapolation and update
  2) Covariance estimate extrapolation and “update”
  3) Filter gain computation

* Rudolf Kalman, RIAS, and Richard C. Bucy, Johns Hopkins Applied Physics Laboratory, 1961. Related developments were made by Weiner (1949), Bode and Shannon (1950), Robbins and Munro (1951), Kiefer and Wolfowitz (1952), Blum (1958), Zadeh and Ragazzini (1952), Carlton and Follin (1956), Berkson (1956), Swerling (1959), Kolmogorov (1962), and Yaglom (1962)
Covariance Estimate (Part 1)

Substitute

$$P_k(-) = \Phi_{k-1}^{-1}P_{k-1}(+)\Phi_{k-1}^T + Q_{k-1}$$

In

$$P_k(+) = \left[ P_{k-1}^{-1}(-) + H_k^TR_k^{-1}H_k \right]^{-1}$$

To obtain the rate of change of the prior covariance estimate

$$\frac{P_k(-) - P_{k-1}(-)}{\Delta t} = F_{k-1}P_{k-1}(-) + P_{k-1}(-)F_{k-1}^T + \frac{Q_{k-1}}{\Delta t} - \frac{K_{k-1}}{\Delta t} H_{k-1}P_{k-1}(-)$$

$$- F_{k-1}K_{k-1}H_{k-1}P_{k-1}(-) - K_{k-1}H_{k-1}P_{k-1}(-)F_{k-1}^T$$

Disturbance and Gain Matrices

Disturbance spectral density matrix

$$Q_{k-1} \approx L(t)Q'_C L^T(t)\Delta t$$

$$\therefore \frac{Q_{k-1}}{\Delta t} \xrightarrow{\Delta t \to 0} L(t)Q'_C(t)L^T(t)$$

Gain matrix

$$K_{k-1} = P_{k-1}(+)H_k^TR_k^{-1}$$

Measurement error covariance matrix

$$R_{k-1} \Delta t = \int_{-\Delta t/2}^{\Delta t/2} R_C \delta(t_{k-1} - \tau) d\tau$$

$$R_{k-1} = \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} R_C \delta(t_{k-1} - \tau) d\tau = \frac{R_C(t_{k-1})}{\Delta t}$$

$$R_{k-1}^{-1} = R_C^{-1}(t_{k-1})\Delta t$$
Covariance Estimate (Part 2)

\[ K_{k-1} = P_{k-1}(+)H_{k-1}^T R_C^{-1} \Delta t \]

\[
\frac{P_k(-) - P_{k-1}(-)}{\Delta t} = F_{k-1} P_{k-1}(-) + P_{k-1}(-) F_{k-1}^T + \frac{Q_{k-1} - P_{k-1}(+)H_{k-1}^T R_C^{-1}H_{k-1} P_{k-1}(-)}{\Delta t} \]

\[
= \begin{bmatrix} F_{k-1} P_{k-1}(+)H_{k-1}^T R_C^{-1}(t_{k-1})H_{k-1} P_{k-1}(-) \\ + P_{k-1}(+)H_{k-1}^T R_C^{-1}(t_{k-1})H_{k-1} P_{k-1}(-) F_{k-1}^T \end{bmatrix} \Delta t
\]

\[ (+) \text{ and } (-) \text{ values coalesce as } \Delta t \to 0 \]

Propagation and update become concurrent

\[ P_{k-1}(+)_{\Delta t \to 0} \to P_{k-1}(-) = P(t_{k-1}) \]

Covariance Estimate (Part 3)

\[ \lim_{\Delta t \to 0} \frac{P_k(-) - P_{k-1}(-)}{\Delta t} = \dot{P}(t) \]

\[ = F(t)P(t) + P(t)F^T(t) + L(t)Q_C(t)L^T(t) - P(t)H^T(t)R_C^{-1}(t)H(t)P(t) \]

\[ \dot{P}(t) = F(t)P(t) + P(t)F^T(t) + L(t)Q_C(t)L^T(t) - K_C(t)H(t)P(t) \]

The continuous-time filter gain matrix is

\[ K_C(t) = P(t)H^T(t)R_C^{-1}(t) \]
Development of Continuous-Time State Estimate

Combine propagation and update equations

\[
\hat{x}_k(+) = \Phi_{k-1} \hat{x}_{k-1}(+) + \Gamma_{k-1} u_{k-1}
\]

\[
\approx \left[ I_n + F(t_{k-1}) \Delta t \right] \hat{x}_{k-1}(+) + G(t_{k-1}) \Delta t u(t_{k-1})
\]

\[
\hat{x}_k(+) = \hat{x}_k(-) + K_k \left[ z_k - H_k \hat{x}_k(-) \right]
\]

\[
\hat{x}_k(+) = \left[ I_n + F(t_{k-1}) \Delta t \right] \hat{x}_{k-1}(+)
\]

\[+G(t_{k-1}) \Delta t u(t_{k-1}) + K_k \left\{ z_k - H_k \left[ I_n + F(t_{k-1}) \Delta t \right] \hat{x}_{k-1}(+) + G(t_{k-1}) \Delta t u(t_{k-1}) \right\}
\]

Continuous-Time State Estimator

In the limit, the difference equation converges to a differential equation

As \( \Delta t \to 0 \), \( t_{k-1} \to t_k \to t \) and

\[
\lim_{\Delta t \to 0} \frac{\hat{x}(t_k) - \hat{x}(t_{k-1})}{\Delta t} = \frac{d\hat{x}(t)}{dt}
\]

State estimator

\[
\hat{x}(t) = F(t)\hat{x}(t) + G(t)u(t) + K_c(t)\left[ z(t) - H(t)\hat{x}(t) \right]
\]

Residual acts as a driving term in the state estimate equation
Summary of the Kalman-Bucy Filter

*State Estimator*

\[
\dot{\hat{x}}(t) = F(t)\hat{x}(t) + G(t)u(t) + K_C(t)\left[ z(t) - H(t)\hat{x}(t) \right]
\]

*Filter Gain Matrix*

\[
K_C(t) = P(t)H^T(t)R_C^{-1}(t)
\]

*Covariance Estimator*

\[
P(t) = F(t)P(t) + P(t)F^T(t) + L(t)Q_C(t)L^T(t) - K_C(t)H(t)P(t),
\]

with \(P(0)\) given.

Block Diagram of the Kalman-Bucy Filter
Second-Order Example
Second-Order Example of Kalman Filter

Rolling motion of an airplane, continuous-time

\[
\begin{bmatrix}
\Delta \dot{\phi}(t) \\
\Delta \dot{\theta}(t)
\end{bmatrix} =
\begin{bmatrix}
L_p & 0 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
\Delta \phi(t) \\
\Delta \theta(t)
\end{bmatrix} +
\begin{bmatrix}
L_{\delta}(t) & 0 \\
0 & 0
\end{bmatrix}\Delta \delta(t) +
\begin{bmatrix}
L_p & 0 \\
0 & 0
\end{bmatrix}\Delta \phi(t)
\]

State estimator

\[
\begin{bmatrix}
\Delta \dot{\phi}(t) \\
\Delta \dot{\theta}(t)
\end{bmatrix} =
\begin{bmatrix}
L_p & 0 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
\Delta \dot{\phi}(t) \\
\Delta \dot{\theta}(t)
\end{bmatrix} +
\begin{bmatrix}
L_{\delta}(t) & 0 \\
0 & 0
\end{bmatrix}\Delta \delta(t) +
\begin{bmatrix}
k_{11}(t) & k_{12}(t) \\
k_{21}(t) & k_{22}(t)
\end{bmatrix}
\begin{bmatrix}
\Delta \phi(t) \\
\Delta \theta(t)
\end{bmatrix} -
\begin{bmatrix}
h_{11} & h_{12} \\
h_{21} & h_{22}
\end{bmatrix}
\begin{bmatrix}
\Delta \dot{\phi}(t) \\
\Delta \dot{\theta}(t)
\end{bmatrix}
\]

Filter gain matrix

\[
\begin{bmatrix}
k_{11}(t) & k_{12}(t) \\
k_{21}(t) & k_{22}(t)
\end{bmatrix} =
\begin{bmatrix}
p_{11}(t) & p_{12}(t) \\
p_{21}(t) & p_{22}(t)
\end{bmatrix}
\begin{bmatrix}
h_{11}(t) & h_{12}(t) \\
h_{21}(t) & h_{22}(t)
\end{bmatrix}
\begin{bmatrix}
r_{11}(t) & 0 \\
0 & r_{22}(t)
\end{bmatrix}
\]

Second-Order Example of Kalman Filter

\[
\begin{bmatrix}
\dot{p}_{11}(t) & \dot{p}_{12}(t) \\
\dot{p}_{21}(t) & \dot{p}_{22}(t)
\end{bmatrix} =
\begin{bmatrix}
L_p & 0 \\
1 & 0
\end{bmatrix}
\begin{bmatrix}
p_{11}(t) & p_{12}(t) \\
p_{21}(t) & p_{22}(t)
\end{bmatrix} +
\begin{bmatrix}
p_{11}(t) & p_{12}(t) \\
p_{21}(t) & p_{22}(t)
\end{bmatrix}
\begin{bmatrix}
L_p & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
p_{11}(t) & p_{12}(t) \\
p_{21}(t) & p_{22}(t)
\end{bmatrix}
\]

\[
-\begin{bmatrix}
k_{11}(t) & k_{12}(t) \\
k_{21}(t) & k_{22}(t)
\end{bmatrix}
\begin{bmatrix}
h_{11}(t) & h_{12}(t) \\
h_{21}(t) & h_{22}(t)
\end{bmatrix}
\begin{bmatrix}
p_{11}(t) & p_{12}(t) \\
p_{21}(t) & p_{22}(t)
\end{bmatrix}
\]
Linear-Optimal Predictor

- $t_k$: Current time, sec
- $t_K$: Future time, sec

**State estimate extrapolation (or propagation)**

\[ \hat{x}(t) = F \hat{x}(t) + Gu(t), \hat{x}(t_k) \text{ from Kalman-Bucy filter} \]

\[ \hat{x}(t_K) = \hat{x}(t_k) + \int_{t_k}^{t_K} \left[ F(\tau)\hat{x}(\tau) + G(\tau)\Delta u(\tau) \right] d\tau \]

**Covariance estimate extrapolation (or propagation)**

\[ \dot{P}(t) = F(t)P(t) + P(t)F^T(t) + L(t)Q'^C \sigma^T L^T(t), \ P(t_k) \text{ from Kalman-Bucy filter} \]

\[ P(t_K) = P(t_k) + \int_{t_k}^{t_K} \left[ F(\tau)P(\tau) + P(\tau)F^T(\tau) + L(\tau)Q'^C(\tau)\sigma^T L^T(\tau) \right] d\tau \]

Duality Between Estimation and Control
Duality Between Linear-Optimal Estimation and Control

Linear-Gaussian Estimator

\[
\dot{x}(t) = F(t)\hat{x}(t) + G(t)u(t) + K_c(t)\left[ z(t) - H(t)\hat{x}(t) \right]
\]

\[
K_c(t) = P(t)H^T(t)R_c^{-1}(t)
\]

\[
\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + L(t)Q_c(t)L^T(t) - P(t)H^T(t)R_c^{-1}(t)H(t)P(t), \quad P(0) \text{ given}
\]

Linear-Quadratic Controller

\[
\dot{x}(t) = F(t)x(t) - G(t)C(t)x(t)
\]

\[
C(t) = R^{-1}(t)G^T(t)S(t)
\]

\[
\dot{S}(t) = -S(t)F(t) - F^T(t)S(t) - Q(t) + S(t)G(t)R^{-1}(t)G^T(t)S(t), \quad S(t_f) \text{ given}
\]

Dual Matrix Definitions for Solution of the Riccati Equation

<table>
<thead>
<tr>
<th>Linear-Quadratic Controller</th>
<th>Linear-Gaussian Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F(t) )</td>
<td>( F^T(t) )</td>
</tr>
<tr>
<td>( G(t) )</td>
<td>( H^T(t) )</td>
</tr>
<tr>
<td>( Q(t) )</td>
<td>( L(t)Q_c(t)L^T(t) )</td>
</tr>
<tr>
<td>( R )</td>
<td>( R_c )</td>
</tr>
<tr>
<td>( S(t) )</td>
<td>( P(t) )</td>
</tr>
<tr>
<td>( S(t_f) )</td>
<td>( P(0) )</td>
</tr>
<tr>
<td>( C(t) )</td>
<td>( K_c(t) )</td>
</tr>
</tbody>
</table>

Controller Riccati equation propagates backward in time

Estimator Riccati equation propagates forward in time
Stability of the Kalman-Bucy Filter

Stochastic Equilibrium of the Constant-Gain Kalman-Bucy Filter

- For linear, time-invariant systems, covariance estimate approaches a non-zero steady state

\[
\dot{P}(t) = FP(t) + P(t)F^T + LQ_C L^T - K_C(t)HP(t) \xrightarrow[\Delta t \to 0]{} 0
\]

- Steady-state covariance matrix is the positive-definite solution to an algebraic Riccati equation

\[
FP_{SS} + P_{SS}F^T + LQ_C L^T - K_CHP_{SS} = 0
\]

- Corresponding filter gain is constant

\[
K_C = P_{SS}H^T R_C^{-1}
\]
Asymptotic Stability of the Constant-Gain Kalman-Bucy Filter

Estimation error
\[ \epsilon(t) = x(t) - \hat{x}_k(t); \quad \dot{\epsilon}(t) = \dot{x}(t) - \dot{\hat{x}}_k(t) \]

Linear, time-invariant state estimator
\[ \hat{x}(t) = F\hat{x}(t) + Gu(t) + K_c[z(t) - H\hat{x}(t)] \]
\[ = (F - K_cH)\hat{x}(t) + Gu(t) + K_cz(t) \]

Constant filter gain matrix
\[ K_c = P_{SS}H^TR_C^{-1} \]

Algebraic Riccati equation
\[ FP_{SS} + P_{SS}F^T + LQ_cL^T - K_cHP_{SS} = 0 \]

Asymptotic Stability of the Steady-State Kalman-Bucy Filter

LTI state equation
\[ \dot{x}(t) = Fx(t) + Gu(t) + Lw(t) \]

Measurement equation
\[ z(t) = Hx(t) + n(t) \]

Estimation error propagation
\[ \dot{\epsilon}(t) = \dot{x}(t) - \dot{\hat{x}}_k(t) \]
\[ = \left[ Fx(t) + Gu(t) + Lw(t) \right] - \left[ F\hat{x}(t) + Gu(t) + K_c[z(t) - H\hat{x}(t)] \right] \]
\[ \dot{\epsilon}(t) = (F - KH)\epsilon(t) + Lw(t) - K_cn(t) \]

Estimator stability governed by \((F - KH)\)
Estimator response driven by \(w(t)\) and \(n(t)\)
**Information Matrix**

*Information matrix and its time rate of change*

\[
\mathbf{J}(t) = \mathbf{P}^{-1}(t) \\
\dot{\mathbf{J}}(t) = \dot{\mathbf{P}}^{-1}(t) = -\mathbf{P}^{-1}(t)\dot{\mathbf{P}}(t)\mathbf{P}^{-1}(t) = -\mathbf{J}(t)\dot{\mathbf{P}}(t)\mathbf{J}(t)
\]

Substitute information matrix in matrix Riccati equation

\[
\mathbf{J}(t) = -\left[ \mathbf{J}(t)\mathbf{F} + \mathbf{F}^T \mathbf{J}(t) + \mathbf{J}(t)\mathbf{L}^T\mathbf{R}^{-1}\mathbf{L} \right]
\]

Steady-state solution for the information matrix

\[
0 = \mathbf{J}_{ss}\mathbf{F} + \mathbf{F}^T \mathbf{J}_{ss} + \mathbf{J}_{ss}\mathbf{L}^T\mathbf{R}^{-1}\mathbf{L} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}
\]

**Lyapunov Function for Estimator Error**

*Lyapunov function for estimation error*

\[
\mathcal{V}[\mathbf{e}(t)] = \mathbf{e}^T(t)\mathbf{J}_{ss}\mathbf{e}(t)
\]

*Homogeneous error dynamics*

\[
\dot{\mathbf{e}}(t) = (\mathbf{F} - \mathbf{K}\mathbf{H})\mathbf{e}(t)
\]

*Time rate of change of the Lyapunov function must be negative to assure stability*

\[
\frac{d\mathcal{V}[\mathbf{e}(t)]}{dt} = 2\mathbf{e}^T(t)\mathbf{J}_{ss}\dot{\mathbf{e}}(t)
\]

\[
= \mathbf{e}^T(t)\left[ \left( \mathbf{J}_{ss}\mathbf{F} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \right) + \left( \mathbf{J}_{ss}\mathbf{F} - \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \right)^T \right]\mathbf{e}(t)
\]

\[
= -\mathbf{e}^T(t)\left[ \mathbf{J}_{ss}\mathbf{L}\mathbf{Q}_c \mathbf{L}^T\mathbf{R}^{-1}\mathbf{L} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \right]\mathbf{e}(t) < 0
\]

\[
\left[ \mathbf{J}_{ss}\mathbf{L}\mathbf{Q}_c \mathbf{L}^T\mathbf{R}^{-1}\mathbf{L} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \right] \text{ is positive definite}
\]
Eigenvalues of the Constant-Gain Kalman-Bucy Filter

LTI state estimation error

\[ \dot{e}(t) = (F - KH)e(t) + Lw(t) - K_Cn(t) \]

Stability indicated by eigenvalues of \((F - KH)\)

\[ |sI_n - (F - KH)| = \Delta_{estimator}(s) = 0 \]

By duality to the LQ regulator, stability of the estimate is guaranteed if the model is correct and

- \((F, H)\): Detectable pair
- \((F, D)\): Stabilizable pair, where \(LWL' = D'D\)

- \(LWL'\): Positive semi-definite matrix
- \(R_C\): Positive definite matrix

Stability of the Estimate vs. Stability of the System

Estimate error is stable even if the system is not

\[ \dot{x} = ax, \quad a = \pm 1, \quad x(0) = 1 \]
\[ \hat{x}(0) = 0.5 \text{ (stable)}, \quad 2 \text{ (unstable)} \]
\[ z = x \]
\[ q = r = 1 \]
Filter Divergence

Could the state estimate diverge from its most likely mean value?

- Yes, if there are
  - Discrepancies in the dynamic model, e.g., nonlinear system with linear model
  - Errors in the assumed covariance matrices
  - Measurement biases
  - Numerical errors in calculation

RMS error = "root-mean-square" error = estimated standard deviation
RSS error = "root-of-sum-of-square" errors = square root of the empirical sum of the squares of the difference between actual and estimated state components

Examples of Filter Divergence
(Schlee, Standish, Toda, 1967)

- Satellite altitude estimation using simplified 1-D model
  - Filter designed assuming constant altitude
  - Drag and gravitational effects change as altitude increases or decreases
Examples of Filter Divergence
(Schlee, Standish, Toda, 1967)

Orbit determination for a space vehicle
Measurements are angle sightings of known terrestrial landmarks
Increased model precision can reduce divergence rate

- **6 - state filter**: Spacecraft position & velocity
- **9 - state filter**: Spacecraft position & velocity, Landmark position

Landmark and Star Tracking

- Landmark and star tracking are functionally similar
- Instrument has narrow field of view
- Star/geographic location catalogs help identify targets
- \( x \) and \( y \) location of landmark or star on focal plane determines angles to the target
Divergence Occurs When Filter Gains are Too Low

1) Dynamic model is incorrect
2) State-error covariance estimate is incorrect
3) Filter gains become too low to properly weight new information

Solutions for Filter Divergence

- Increase “process noise” used for estimator design (i.e., assumed disturbance covariance)
- Improve system modeling, e.g.,
  - Estimate measurement bias and scale factor
  - Include higher-order terms
  - Model nonlinearity
- Use higher precision or square-root filtering
- Adapt estimator to changing conditions
Recall Effects of Assumed “Process Noise” and Measurement Error

Adding “Process Noise” to Eliminate Divergence*

• Satellite orbit determination
  – Aerodynamic drag produced unmodeled bias
  – Optimal filter did not estimate bias

• “Process noise” increased for filter design
  – Divergence was contained

* Fitzgerald, 1971
Square Root Filtering

Improved precision by reducing condition number

Define

\[ P(t) = S(t)S^T(t) \]

where

\[ S(t) : \text{ Lower triangular square root of } P(t) \]

Use Cholesky decomposition to compute \( S(t) \) [see text]

Riccati equation

\[
\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + L(t)Q'_C(t)L^T(t) - P(t)H^T(t)R_C^{-1}(t)H(t)P(t)
\]

Substitute

\[
S(t)S^T(t) + S(t)S^T(t) = F(t)S(t)S^T(t) + S(t)S^T(t)F^T(t) + L(t)Q'_C(t)L^T(t) - S(t)S^T(t)H^T(t)R_C^{-1}(t)H(t)S(t)S^T(t)
\]
Matrix Decomposition Examples

Cholesky Decomposition

\[
\begin{pmatrix}
4 & 12 & -16 \\
12 & 37 & -43 \\
-16 & -43 & 98
\end{pmatrix}
= \begin{pmatrix}
2 & 6 & -8 \\
6 & 1 & 3 \\
-8 & 5 & 3
\end{pmatrix}
\]

UDU^T (or LDL^T) Decomposition

\[
\begin{pmatrix}
4 & 12 & -16 \\
12 & 37 & -43 \\
-16 & -43 & 98
\end{pmatrix}
= \begin{pmatrix}
1 & 3 & -4 \\
3 & 1 & 9 \\
-4 & 5 & 1
\end{pmatrix}
\]

Square-Root Filtering

Pre-multiply by \( S^{-1}; \) post-multiply by \( S^{-T} \)

\[
S^{-1} \hat{S} + \hat{S}^T S^{-T} = S^{-1} F S + S^T F^T S^{-T} + S^{-1} L Q'_{C} L^T S^{-T} - S^T H^T R_{C}^{-1} H S
\]

\[\cong M(t) = M_{LT}(t) + M_{UT}(t)\]

Elements of \( M_{LT}(t) \)

\[
\begin{cases}
  m_{ij}, & i > j \\
  m_{ij}/2, & i = j \\
  0, & i < j
\end{cases}
\]

\( M_{LT}(t) \) : Lower triangular portion of \( M(t) \)
\( M_{UT}(t) \) : Upper triangular portion of \( M(t) \)

Because \( S \) is lower triangular
\( \hat{S} \) and \( S^{-1} \) are lower triangular
Square-Root Filtering

Hence, the Riccati equation becomes

\[ \dot{S}(t) = S(t)M_{LT}(t), \quad S(0)S^T(0) = P(0) > 0 \]

Estimator gain matrix

\[ K_C(t) = S(t)S^T(t)H^T(t)R_C^{-1}(t), \quad S(0)S^T(0) = P(0) > 0 \]

Square root required to define $S(0)$ but not for calculation of $S(t)$

Kalman Filter for Correlated Disturbance Inputs and Measurement Noise*

Disturbance process produces error in measurements

Correlation expressed by expected value

\[ E\left[ \begin{bmatrix} w(t) \\ n(t) \end{bmatrix} \begin{bmatrix} w(\tau) & n(\tau) \end{bmatrix} \right] = \begin{bmatrix} Q_C(t) & M_C(t) \\ M_C^T(t) & R_C(t) \end{bmatrix} \delta(t - \tau) \]

Riccati equation

\[ \dot{P}(t) = F(t)P(t) + P(t)F^T(t) + L(t)Q_C(t)L^T(t) - \left[ P(t)H^T(t) + L(t)M_C(t) \right]R_C^{-1}(t)\left[ H(t)P(t) + M_C^T(t)L^T(t) \right] \]

Estimator gain matrix

\[ K_C(t) = \left[ P(t)H^T(t) + L(t)M_C(t) \right]R_C^{-1}(t) \]

* Bryson and Johansen
Next Time:
Nonlinear State Estimation

Supplemental Material
Program for Example of Kalman Filter Estimate Error Stability

% Kalman-Bucy Filter Estimate Stability
% Copyright by Robert Stengel. All rights reserved.
% 4/11/2010

clear

% First-Order Stable and Unstable Systems
StableSys = ss(-1,1,1,0);
UnstableSys = ss(1,1,1,0);

[Kstable,Lstable,Pstable] = kalman(StableSys,1,1,0);
[Kunstable,Lunstable,Punstable] = kalman(UnstableSys,1,1,0);

StableSysEst = ss((-1-Lstable),1,1,0);
UnstableSysEst = ss((1-Lunstable),1,1,0);

% Response

t = [0:0.01:6];
y1,t1,x1 = initial(StableSys,1,t);
y2,t2,x2 = initial(UnstableSys,1,t);
y3,t3,x3 = initial(StableSysEst,0.5,t);
y4,t4,x4 = initial(UnstableSysEst,-1,t);

figure
y3 = interp1(t3,y3,t1);
y4 = interp1(t4,y4,t2);
plot(t1,y1,t2,y2,t1,(y1-y3),t2,(y2-y4)),grid, axis([0 6 0 4])