The Kalman Filter

03-04-2014

sol^{n} :

$$\min J = \frac{1}{2} x^T(t_1) P_1 x(t_1) + \frac{1}{2}\int_{t_0}^{t_1} \left[x^T R_1 x + u^T R_2 u\right] dt$$

s.t. $\dot{x} = Ax + Bu$

Sol^{n} : $u(t) = -\overbrace{R_2^{-1} B^T P(t)}^{F^*(t)} x(t)$

where $\boxed{\begin{aligned} -\dot{P} &= R_1 - PBR_2^{-1}B^T P + PA + A^T P \\ \text{with } & P(t_1) = P_1 \end{aligned}}$

Also recall that

$$J^*(t, x(t)) = x^T(t) P(t) x(t)$$

Lemma: Consider a matrix diff. eqn:

$$-\dot{\tilde{P}}(t) = R_1 + F^T(t) R_2 F(t) + \tilde{P}(t)\left[A - BF(t)\right] + \left[A - BF(t)\right]^T \tilde{P}(t)$$

with $\tilde{P}(t_1) = P_1$ (with $R_1, P_1 \geq 0$, $R_2 > 0$)

Then:

1> $\tilde{P}(t) \geq P(t)$ for $t_0 \leq t \leq t_1$

2> $\tilde{P}(t) = P(t)$ if $F(t) = R_2^{-1} B^T P(t)$

Proof: Easy: recall $J(t, x(t)) = x^T P(t) x(t)$ is minimized by $u(t) = -R_2^{-1} B P(t)$

So $x^T \tilde{P}(t) x(t) \geq x^T P(t) x(t)$ $\forall x(t)$.

Stochastic Processes & Linear Systems

Let $v(t)$ be a vector-valued stochastic process:

<u>Mean</u>: $m(t) = E\{v(t)\}$

<u>Covariance Matrix</u>: $R_v(t_1, t_2)$

$$= E\{[v(t_1) - m(t_1)][v(t_2) - m(t_2)]^T\}$$

<u>Variance Matrix</u>: $R_v(t,t) = Q(t)$

<u>2nd Order joint moment matrix</u>:

$$C_v(t_1, t_2) = E\{v(t_1)\, v^T(t_2)\}$$

If $m(t) = 0 \quad \forall t$, $R_v(t_1, t_2) = C_v(t_1, t_2)$

$$C_v(t_1, t_2) = \begin{bmatrix} E[v_1(t_1) v_1(t_2)] & \cdots & E[v_1(t_1) v_m(t_2)] \\ \vdots & & \\ E[v_m(t_1) v_1(t_2)] & \cdots & E[v_m(t_1) v_m(t_2)] \end{bmatrix}$$

<u>FACT</u>: Let $W(t)$ be square symmetric:

$$E\{v^T(t) W(t) v(t)\} = tr[W(t) C_v(t,t)]$$

$$E\{v^T W v\} = E\left\{\sum_{i,j=1}^{n} v_i W_{ij} v_j\right\}$$

$$= \sum_{i,j=1}^{n} W_{ij} E\{v_i v_j\} = \sum_{i,j=1}^{n} W_{ij}\, C_{v,ij}(t,t)$$

$$= tr[W(t) C_v(t,t)]$$

White Noise : Zero mean with

$$R_W(t_1,t_2) = V(t_1)\,\delta(t_2-t_1)$$
$$V(t) \geq 0$$

Some Integration Formulae : for

~~$w(t) \to$ white noise with~~ $V(t)$

a) $E\left\{ \int_{t_1}^{t_2} A(t)\, w(t)\, dt \right\} = 0$

b) $E\left\{ \left[\int_{t_1}^{t_2} A_1(t)\, w(t)\, dt \right]^T W \left[\int_{t_3}^{t_4} A_2(t')\, w(t')\, dt' \right] \right\}$

$$= \int_I \operatorname{tr}\left[V(t)\, A_1^T(t)\, W\, A_2(t) \right] dt$$

$I = [t_1, t_2] \cap [t_3, t_4]$, W any weight matrix.

c) $E\left\{ \left[\int_{t_1}^{t_2} A_1(t)\, w(t)\, dt \right] \left[\int_{t_3}^{t_4} A_2(t')\, w(t')\, dt' \right]^T \right\}$

$$= \int_I A_1(t)\, V(t)\, A_2^T(t)\, dt$$

Linear system with white noise input:

FACT: Let $\dot{x}(t) = Ax + Bw(t)$, $x(t_0) = x_0$ | $w(t) \sim V(t)$, $E(w(t)) = 0$

x_0 is a stochastic variable ind. of $w(t)$ with mean m_0 and variance $Q_0 = E\left[(x_0 - m_0)(x_0 - m_0)^T\right]$

Then the mean of $x(t)$

$$m_x(t) = \underbrace{\phi(t, t_0)}_{e^{A(t-t_0)}}\, m_0$$

$$R_x(t_1,t_2) = \Phi(t_1,t_0)\, Q_0\, \Phi^T(t_2,t_0) + \int_{t_0}^{\min(t_1,t_2)} \Phi(t_1,\tau)\, B\, V(\tau)\, B^T \Phi^T(t_2,\tau)\, d\tau$$

$$Q(t) = R_x(t,t)$$

$$\dot{Q} = AQ + QA^T + BVB^T$$

$$Q(t_0) = Q_0$$

Hints for Proof:

$$x(t) = \Phi(t,t_0)\, x_0 + \int_{t_0}^{t} \Phi(t,\tau)\, B\, w(\tau)\, d\tau$$

$$m_x(t) = \Phi(t,t_0)\, E(x_0) + 0.$$

Derivation hints for $C_x(t,t_2) \rightarrow R_x(t_1,t_2)$ similar.

Now, $E\{x(t_1)\, x^T(t_2)\} =$ derive (Exercise)

$$= \Phi(t_1,t_0)\, E(x_0 x_0^T)\, \Phi^T(t_2,t_0) + E\left\{ \left[\int_{t_0}^{t_1} \Phi(t_1,\tau)\, B\, w(\tau)\, d\tau\right] \left[\int_{t_0}^{t_2} \Phi(t_2,\tau)\, B\, w(\tau)\, d\tau\right]^T \right\}$$

$\longrightarrow$ then use FACT above.

$$= \Phi(t_1,t_0)\, C_x(t_0,t_0)\, \Phi^T(t_2,t_0) + \int_{t_0}^{\min(t_1,t_2)} \Phi(t_1,\tau)\, B\, V\, B^T \Phi^T(t_2,\tau)\, d\tau$$

$$\underbrace{Q(t) = R_x(t,t)}_{F}$$

$$\dot{Q} = \frac{d}{dt}\left[e^{A(t-t_0)} Q_0 e^{A^T(t-t_0)} + \int_{t_0}^{t} e^{A(t-\tau)} B V(\tau) B^T e^{A^T(t-\tau)} d\tau \right]$$

$$= A\left(F \right) + \left(F \right) A^T$$

$$+ \int_{t_0}^{t} \frac{d}{dt}\left[\quad \right] d\tau + \left[\quad \right]_{\tau = t}$$

$$= A\left(F \right) + \left(F \right) A^T + A\left[I \right]$$

$$+ \left[I \right] A + I B V(t) B^T I$$

$$= AQ + QA^T + BVB^T$$

The Optimal Observer

System : $\dot{x}(t) = Ax(t) + Bu(t) + w_1(t)$

$$y(t) = C x(t) + w_2(t) \qquad (1)$$

$w_1(t) \rightarrow$ state excitation noise

$w_2(t) \rightarrow$ observation / measurement noise

$[w_1^T(t) \; w_2^T(t)] \sim$ white noise with

$$V(t) = \begin{bmatrix} V_1(t) & V_{12}(t) \\ V_{12}^T(t) & V_2(t) \end{bmatrix}$$

i.e. $E\left\{ \begin{bmatrix} w_1(t_1) \\ w_2(t_1) \end{bmatrix} \left[w_1^T(t_2) \;\; w_2^T(t_2) \right] \right\} = V(t_1)\delta(t_1 - t_2)$

We assume 1) $V_{12}(t) = 0 \equiv \omega_1(t)$ & $\omega_2(t)$ are uncorrelated.

2) $V_2(t) > 0 \quad \forall\, t \geq t_0$

<u>Denote</u>: $E\{x(t_0)\} = \bar{x}_0$

$$E\{[x(t_0) - \bar{x}_0][x(t_0) - \bar{x}_0]^T\} = Q_0$$

<u>Observer Equations</u>:

$$\dot{\hat{x}} = A\hat{x} + Bu + K(t)[y - C\hat{x}] \quad (2)$$

<u>Error</u>: $e(t) = x(t) - \hat{x}(t)$

<u>Mean Square Error</u>: $E\{e^T(t)W(t)e(t)\}$ with $W(t) > 0$ weighing matrix

<u>Optimal Observer Problem</u>:

Choose 1) $K(\tau) \quad t_0 \leq \tau \leq t$

2) $\hat{x}(t_0)$

to minimize $E\{e^T(t)\,W(t)\,e(t)\}$

(Kalman & Bucy, 1961)

<u>Time reversal Lemma</u>:

Consider $\dfrac{dx}{dt} = f(t, x(t)) \qquad t \geq t_0$

$x(t_0) = x_0$

$$-\frac{dy}{dt} = f(t^* - t, y(t)) \qquad t \leq t_1$$

$y(t_1) = y_1$

$$\overbrace{t_0 \longrightarrow t^*}^{t_1}$$

where $t_0 < t_1$ & $\underline{t^* = t_0 + t_1}$

Then if $x_0 = y_1$ then

$$x(t) = y(t^* - t) \qquad t \geq t_0$$
$$y(t) = x(t^* - t) \qquad t \leq t_1$$

From ① & ②,

$$\dot{e}(t) = (A - KC)e + [I \;\; -K(t)]\begin{bmatrix} \omega_1(t) \\ \omega_2(t) \end{bmatrix}$$

$$e(t_0) = e_0 = x(t_0) - \hat{x}(t_0)$$

Let $E\{e(t)\} = \bar{e}(t)$

$$E\{[e(t) - \bar{e}(t)][e(t) - \bar{e}(t)]^T\} = \tilde{Q}(t)$$

$$\Rightarrow C_e(t,t) = E(e(t)e^T(t)) = \bar{e}(t)\bar{e}^T(t) + \tilde{Q}(t)$$

Then: the mse:

$$E\{e^T(t)W(t)e(t)\} = tr[W(t)C_v(t,t)]$$

$$= tr[W(t)\{\bar{e}(t)\bar{e}^T(t) + \tilde{Q}\}]$$

$$= \bar{e}^T(t)W(t)\bar{e}(t) + \underline{tr(\tilde{Q}W(t))} \quad ③$$

First term is minimized if $\bar{e}(t) = 0$

Now $\dot{\bar{e}}(t) = [A - K(t)C]\bar{e}(t) \qquad t \geq t_0$

Hence choose $\bar{e}(t_0) = 0$, then $\bar{e}(t) = 0$

We can make $\bar{e}(t_0)=0$ by choosing

$$\hat{x}(t_0) = \bar{x}_0$$

Second term in (3) is ind. of $\bar{e}(t)$. Hence independent minimization:

From FACT above:

$$\dot{\tilde{Q}}(t) = [A-K(t)C]\tilde{Q} + \tilde{Q}[A-K(t)C]^T + [I \quad -K(t)]\begin{bmatrix} V_1(t) & 0 \\ 0 & V_2(t)\end{bmatrix}\begin{bmatrix} I \\ -K^T(t)\end{bmatrix} \tag{*}$$

with $\tilde{Q}(t_0) = Q_0$.

Now introduce the time reversed version of (*) — variable $\tilde{P}(t)$

$$-\dot{\tilde{P}}(t) = [A-K(t^*-t)C]\tilde{P}(t) + \tilde{P}[A-K(t^*-t)C]^T + V_1(t^*-t) + V_1(t^*-t) + K(t^*-t)V_2(t^*-t)K^T(t^*-t) \tag{**}$$

where $t^* = t_0 + t_1$, $t_1 > t_0$

<u>Let</u>: $\tilde{P}(t_1) = Q_0$

Then from the time reversal lemma:

$$\tilde{Q}(t) = \tilde{P}(t^*-t) \qquad t \leq t_1$$

Now from LQR theory (last lemma), $\tilde{P}(t)$ is minimized if

$$K^o(t^*-\tau) = V_2^{-1}(t^*-\tau)\,C\,P(\tau)$$
$$t \leq \tau \leq t_1$$

This $P(t)$ can be solved by putting $K^o(t^*-\tau)$ in (**):

$$-\dot{P}(t) = V_1(t^*-t) - P(t)C^T V_2^{-1}(t^*-t)\,C\,P(t) + P(t)A^T + AP \qquad (1)$$

with $P(t_1) = Q_0$

Then: $P(t) \leq \tilde{P}(t) \quad t \leq t_1$

By reversing the time again in (1),

$$\tilde{Q}(t) \geq Q(t) \qquad t \geq t_0$$

by choosing $K^o(\tau) = Q(\tau)C^T V_2^{-1}(\tau)$, $\tau \geq t_0$

where $Q(\tau)$ satisfies:

$$\dot{Q}(t) = V_1(t) - Q(t)C^T V_2^{-1}(t)\,C\,Q(t) + Q(t)A^T + AQ(t)$$

with $Q(t_0) = Q_0$

But $Q(t) \leq \tilde{Q}(t)$

$$\Rightarrow tr[Q(t)W(t)] \leq tr[\tilde{Q}(t)W(t)]$$

for any $W(t) > 0$.

Hence $K^0(t)$ of (2) optimizes the observer i.e. minimizes $E\{e^T(t)\,W(t)\,e(t)\}$ for every time $t \geq t_0$.

At the min: $E\{e^T(t)W(t)e(t)\} = tr[Q(t)W(t)]$
& $E[e(t)e^T(t)] = Q(t)$

<u>Steady state Sol$^{\underline{n}}$: for LTI</u>

$$\dot{x} = Ax(t) + Bu(t) + G\,w_3(t)$$
$$y(t) = Cx(t) + w_2(t)$$

$w_3 \rightarrow$ white noise of intensity V_3
$w_2 \longrightarrow V_2$
$V_3 > 0$, $V_2 > 0$, $Q_0 \geq 0$

▷ If $\{A, G, C\}$ is both stabilizable & detectable the the sol$^{\underline{n}}$ of

$$\begin{cases} \dot{Q}(t) = AQ + QA^T + \underbrace{GV_3G^T}_{V_1 \text{ before}} - Q(t)C^TV_2^{-1}CQ(t) \\ \text{with } Q(t_0) = Q_0 \end{cases}$$

converges to a steady-state value $\overline{Q}$ as $t_0 \rightarrow -\infty$ for every $Q_0 \geq 0$.

then

2) $\bar{Q}$ is the sol^{n} (unique non-negative definite symmetric) of the ARE:

$$0 = AQ + QA^T + G V_3 G^T - QC^T V_2^{-1} C Q$$

3) If $\bar{Q}$ exists, it is +ve def. iff system is completely controllable.

4) If $\bar{Q}$ exists, the steady-state optimal observer

$$\dot{\hat{x}} = A\hat{x} + \bar{K}[y - C\hat{x}]$$

where

$$\bar{K} = \bar{Q} C^T V_2^{-1}$$

is asymp. stable iff system is detectable + stabilizable.

5) If system is detectable + stabilizable

the steady-state optimal observer minimizes:

$$\lim_{t_0 \to -\infty} E\{e^T(t) W e(t)\}$$

for all $Q_0 \geq 0$. For the steady-state optimal observer:

$$\lim_{t_0 \to -\infty} E\{e^T(t) W e(t)\} = \mathrm{tr}[\bar{Q} W]$$

Example: Position Servo

$$\begin{bmatrix} & \\ 0 & -\alpha \end{bmatrix} x + \begin{bmatrix} 0 \\ \kappa \end{bmatrix} \mu(t) + \begin{bmatrix} 0 \\ \nu \end{bmatrix} \tau_d(t)$$

modelled as white noise ← { disturbing high freq torque

$$\eta(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) + \nu_m(t)$$

Intensity of $\left.\begin{matrix} \tau_d(t) \longrightarrow V_d \\ \nu_m(t) \longrightarrow V_m \end{matrix}\right\}$ constant scalar

The variance Riccati Eqn:

$$\dot{Q}(t) = \begin{bmatrix} 0 & 1 \\ 0 & -\alpha \end{bmatrix} Q(t) + Q(t) \begin{bmatrix} 0 & 0 \\ 1 & -\alpha \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \nu^2 V_d \end{bmatrix} - Q(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{V_m} \begin{bmatrix} 1 & 0 \end{bmatrix} Q(t)$$

Let $Q(t) = \begin{bmatrix} q_{11}(t) & q_{12}(t) \\ q_{12}(t) & q_{22}(t) \end{bmatrix}$

$$\dot{q}_{11}(t) = 2q_{12}(t) - \frac{1}{V_m} q_{11}^2(t)$$
$$\dot{q}_{12}(t) = q_{22}(t) - \alpha q_{12}(t) - \frac{1}{V_m} q_{11}(t) q_{12}(t)$$
$$\dot{q}_{22}(t) = -2\alpha q_{22}(t) + \nu^2 V_d - \frac{1}{V_m} q_{12}^2(t)$$

By putting $\dot{Q}(t) = 0$, one can solve for $\bar{Q}$.

$$\bar{K} = \begin{bmatrix} -\alpha + \sqrt{\alpha^2 + 2\beta} \\ \alpha^2 + \beta - \alpha\sqrt{\alpha^2 + 2\beta} \end{bmatrix}. \qquad \beta = \nu\sqrt{\frac{V_d}{V_m}}$$

$\ldots) = s^2 + s\sqrt{\alpha^2+2\beta} + \beta$

Poles of Optimal Observer: $\frac{1}{2}\left(-\sqrt{\alpha^2+2\beta} \pm \sqrt{\alpha^2-2\beta}\right)$

Observation: As $\beta \uparrow$ poles move left.

$\left(\beta \propto \frac{V_d}{V_m} \equiv \frac{\text{signal}}{\text{noise}}\right)$

Faster state reconstruction but more sensitive to observation noise.

Equivalent Results for LQR for LTI
(should have been done after LQR)

Consider: $\dot{x}(t) = Ax(t) + Bu(t)$
$z(t) = Dx(t)$

$$\min \; J = x^T(t_1)P_1 x(t_1) + \int_{t_0}^{t_1}\left\{z^T(t)R_3 z(t) + u^T(t)R_2 u(t)\right\}dt$$

$R_3 > 0,\; R_2 > 0,\; P_1 \geq 0$

$$\left[\text{Note } J = \left[\; J_{t_1} + \int_{t_0}^{t_1}\left\{x^T\underbrace{[D^T R_3 D]}_{R_1}x + u^T R_2 u\right\}dt\right]\right]$$

Diff. Riccati Eqn:

$$(*)\begin{cases} -\dot{P}(t) = D^T R_3 D - P(t)BR_2^{-1}B^TP(t) + A^TP(t) + P(t)A \\ \text{with} \quad P(t_1) = P_1 \end{cases}$$

1) If system is stabilizable + detectable the sol$^{\underline{n}}$ of (*) approaches unique $\bar{P}$ as

... every $P_1 \geq 0$

2) If (stab + Det.) $\rightarrow$ $\bar{P}$ is the unique non-negative definite symmetric sol. of the ARE:

$$0 = D^T R_3 D - P B R_2^{-1} B^T P + A^T P + PA$$

3) if $\bar{P}$ exists, $\bar{P} > 0 \Leftrightarrow$ system is completely observable

4) if $\bar{P}$ exists, steady state control law:

$$u(t) = -\bar{F} x(t)$$

$$\bar{F} = R_2^{-1} B^T \bar{P}$$

is asymp. stable iff system is stab + detec.

5) If stab + detec., the steady state control law minimizes:

$$\lim_{t_1 \to \infty} \left\{ x^T(t_1) P_1 x(t_1) + \underbrace{\int_{t_0}^{t_1} \left\{ z^T(t) R_3 z(t) + u^T R_2 u \right\} dt}_{(\ast)} \right\}$$

for all $P_1 \geq 0$. Then

$$(\ast) = x^T(t_0) \bar{P} x(t_0)$$