“State Estimation Using the Kalman Filter”

Thomas F. Edgar
Department of Chemical Engineering
University of Texas
Austin, TX 78712

Outline

- Introduction
- Basic Statistics for Linear Dynamic Systems
- State Estimation
- Kalman Filter – Algorithm and Properties

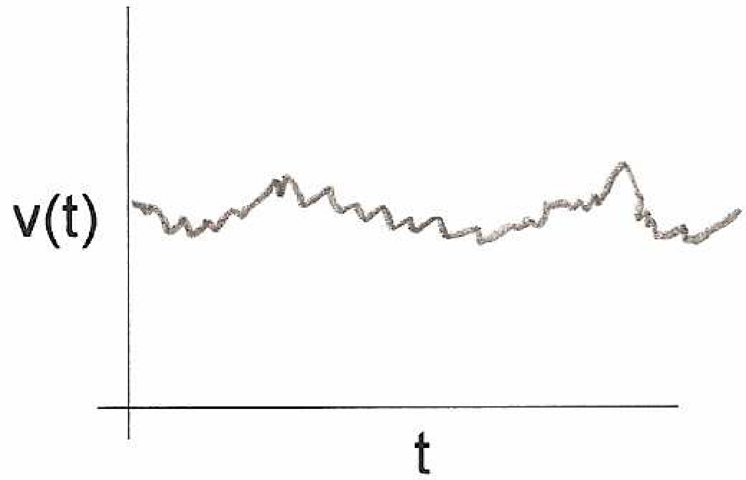
Control with Limited/Noisy Measurements

(1) Some variables may not be measurable in real time

(2) Noise in the instruments and in the process may give erroneous data for control purposes

Solution: Use Kalman Filter

Random Variables



Ex

turbulent flow,
temperature sensor
in boiling liquid

mean (expected) value of r.v. :

$$\bar{x} = E[x] \triangleq \int_{-\infty}^{\infty} x p(x) dx$$

↑ expected value operator

p(x) probability density function

Definition of Variance and Covariance

$$\mathrm{var}(x) \triangleq E\left[(x-\bar{x})^2\right] \triangleq \overline{(x-\bar{x})(x-\bar{x})}$$

$$\int_{-\infty}^{\infty} (x-\bar{x})^2 \; p(x)\; dx$$

$$\left[\mathrm{var}(x)\right]^{\frac{1}{2}} = \sigma \text{ (standard deviation)}$$

for two random variables, x and y, the degree of their dependence is indicated by covariance $= \mathrm{cov}(x,y) \triangleq E\left[(x-\bar{x})(y-\bar{y})\right]$

$$\mathrm{cov}(x,y) = E[xy] - \bar{x}\;\bar{y}$$

$$\mathrm{cov}(x,x) = \mathrm{var}\, x = E\left[x^2\right] - (\bar{x})^2$$

Correlation Coefficient

$$\rho(x,y) \underset{=}{\Delta} \frac{\text{cov}(x,y)}{\sqrt{\text{var}(x)\text{var}(y)}}$$

$$-1 \le \rho \le +1$$

$$|\rho| \to 1$$

highly correlated

$$|\rho| \to 0$$

uncorrelated

extension to multivariable case:

covariance matrix $P_{ij} = \text{cov}(x_i, x_j)$

$$\underline{\underline{P}} \underset{=}{\Delta} E\left[(x-\bar{x})(x-\bar{x})^T\right]$$

$$\underline{\underline{P}} = \underline{\underline{P}}^T \quad \underline{\underline{P}} \ge 0 \quad p_{ii} = \text{var}\, x_i$$

State Estimation

Object: Using data (which is filtered), reconstruct values for unmeasured state variables

Definitions: $\text{mean} = \bar{x} = \frac{1}{N}\sum^{N} x_i$

variance $\sigma^2 = \sum (x_i - \bar{x})^2$

σ^2 large, lots of scatter! single data pt. is unreliable

Example: 2 measurements of equal reliability

$$(x_p, x_r) \rightarrow x_f \qquad \frac{x_p + x_r}{2} = x_f$$

more generally, $\underline{x}_p$ $\underline{x}_r$ data pt. "*p*"
data pt. "*r*"
(actually data vectors)

$$\underline{e}_p = \underline{x}_p - \underline{\bar{x}}$$

$$\underline{e}_r = \underline{x}_r - \underline{\bar{x}}$$

State Estimation (cont'd)

error vectors $\rightarrow$ error covariance matrices

$$\underline{\underline{P}} = \overline{e_p e_p^T} = \begin{bmatrix} \overline{e_{p_1} e_{p_1}} & \overline{e_{p_1} e_{p_2}} & \cdots & \\ & \overline{e_{p_2} e_{p_2}} & \ddots & \\ & & & \overline{e_{p_n} e_{p_n}} \end{bmatrix}$$

Similarly, $\underline{\underline{R}} = \overline{e_r e_r^T}$

combine $\underline{x}_p$ and $\underline{x}_r$ to find $\underline{x}_f$

$$\underline{x}_f = (\underline{\underline{I}} - \underline{\underline{F}})\underline{x}_p + \underline{\underline{F}}(\underline{x}_r) = \underline{x}_p + \underline{\underline{F}}(\underline{x}_r - \underline{x}_p)$$

$$\underline{e}_f = (\underline{\underline{I}} - \underline{\underline{F}})\,\underline{e}_p + \underline{\underline{F}}\,\underline{e}_r$$

$\underline{\underline{F}}$: weighting matrix

The covariance matric of $\underline{e}_f$ is

$$\underline{\underline{H}} = \underline{\underline{P}} - \underline{\underline{F}}\,\underline{\underline{P}} - \underline{\underline{P}}\,\underline{\underline{F}}^T + \underline{\underline{F}}\,(\underline{\underline{P}} + \underline{\underline{R}})\,\underline{\underline{F}}^T$$

(no correlation between $\underline{e}_r$ and $\underline{e}_p$)

State Estimation (cont'd)

Select F so that $\sum_i {e_{f_i}}^2$ is a minimum

(least squares estimate or minimum variance estimate)

$$F^{opt} = P(P+R)^{-1}$$

scalar example:

(*a*) $\begin{matrix} P = q^2 \\ R = q^2 \end{matrix} \qquad F = \frac{1}{2}$

(b) $P >> R \qquad F = 1 \qquad$ (select x_r)

State Estimation (cont'd)

For dynamic systems, we have 2 sources of information:

(x_p) (1) state equation (and previous state estimates)

(x_r) (2) new measurements at time step k
(but # measurements $\leq$ # states)

$$x(k) = A(k\text{-}1)\, x(k\text{-}1) + G(k\text{-}1)\, w(k\text{-}1)$$

linear dynamic system; w = process noise

state variables x_{nx1}

State Estimation (cont'd)

$$y(k) = Cx(k) + v(k)$$

$y_{\ell x1}$ $\quad \ell \leq n$ $\quad v(k)$: instrument noise

y: measured variables

We wish to update $\hat{x}(k)$ (the estimates of the states) from inaccurate or unknown initial conditions on $x(k)$, measurements corrupted by noise.

<u>The Kalman filter eqn:</u>

$$\hat{x}(k) = A(k-1)\hat{x}(k-1)$$

$$+ K(k)\,[y(k) - \underbrace{CA\,(k-1)\,\hat{x}\,(k-1)}_{\text{"}\hat{y}\text{"}}]$$

$y - \hat{y}$: difference between measurement and estimate

n.b. if we can't measure $x_i\,(k)$, then $x_i\,(0)$ is unknown

State Estimation (cont'd)

Define covariance matrices

Q_{nxn} : $w(k)$ process noise vector (white)

$R_{\ell x \ell}$: $v(k)$ instrument noise vector (white)

Q, R usually diagonal matrices that can be "tuned" (variances needed)

w, v are uncorrelated white noises (characteristics defined by mean, variance)

markov sequence $p(w(k+1) / w(k)) = p(w(k+1))$

define state error covariance

$$P = \left(x_p - x\right)\left(x_p - x\right)^T$$

x: true value

State estimate covariance

$$H = (\hat{x} - x)(\hat{x} - x)^T$$

State Estimation (cont'd)

given $\hat{x}(k-1)$, an unbiased estimate

(min. variance) is

$$x_p(k) = A(k-1)\hat{x}(k-1)$$

covariance matrix:

$$P(k) = \overline{\left(x_p - x_k\right)\left(x_p - x_k\right)^T}$$

$$= \overline{\left[A(k-1)\hat{x}(k-1) - A(k-1)x(k-1) - G(k-1)w(k-1)\right] \bullet}$$

$$\overline{[\quad]^T}$$

$$= A(k-1)H(k-1)A(k-1)^T + G(k-1)Q(k-1)G(k-1)^T$$

State Estimation (cont'd)

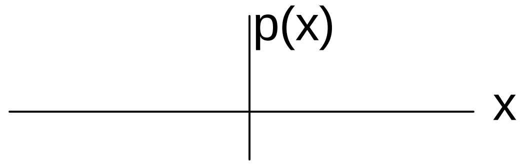
adjust measurement eqn. (to obtain square system):

$$y_1(k) = C_1 x(k) + v_1(k) \qquad y = Cx + v$$

$$nx1 \qquad nxn \qquad nx1 \qquad \ell x1 \quad \ell xn \quad \ell x1$$

$$C_1 = \begin{bmatrix} C \\ \hline D \end{bmatrix}$$ D: dummy matrix so C_1 is square

$\mathrm{cov}(v_1)$ $\quad R_1 = \begin{bmatrix} R & \mathrm{O} \\ \hline O & \gamma I \end{bmatrix}$ $\quad \gamma \to \infty$ (bad or missing information *on* $n-\ell$ *components*)



State Estimation (cont'd)

Inverting,

$$x(k) = C_1^{-1}[y_1(k) - v_1(k)]$$

<u>estimate</u> $x_r(k) = C_1^{-1} y_1(k)$

$$\text{cov}\left[x_r - x(k)\right] = C_1^{-1} R_1(k) (C_1^{-1})^T$$

<u>Ex</u> $C = [\,1 \quad 0\,]$

select $D = [\,0 \quad 1\,]$ to make H_1 invertible

R = 1 $R_1 = \begin{bmatrix} 1 & 0 \\ 0 & \text{N} \end{bmatrix}$ N large

$R_1^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \text{N}^{-1} \end{bmatrix}$ note $N^{-1} \to 0$

potential for roundoff error

State Estimation (cont'd)

Combine $\underbrace{x_r \text{ and } x_p}_{\hat{x}}$ to minimize variance

$$\hat{x}(k) = x_p(k) + F(k)\left[x_r(k) - x_p(k)\right]$$

$$(F^{opt} = P(P+R)^{-1}$$

from earlier analysis)

$$\hat{x}(k) = A(k-1)\hat{x}(k-1) + F(k)[C_1^{-1}y_1(k) - A(k-1)\hat{x}(k-1)]$$

actually, $F^{opt}(k) = P(k)[P(k) + C_1^{-1}R_1(k)(C_1^{-1})^T]^{-1}$

$$F^{opt}(k) = P(k)[C_1^T R_1^{-1}(k) C_1 P(k) + I]^{-1} C_1^T R_1(k)^{-1} C_1$$

State Estimation (cont'd)

examine $C_1^T R_1^{-1} C_1 = \begin{bmatrix} C^T & D^T \end{bmatrix} \begin{bmatrix} R^{-1} & 0 \\ 0 & \gamma^{-1} I \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix}$

$$= C^T R^{-1} C$$

covariance of composite estimate becomes

$$H(k) = [P(k) C^T R^{-1}(k) C + I]^{-1} P(k)$$

$$\uparrow$$

nxn matrix inverse

By rearrangement and matrix identity (reduce size of inverse to be computed)

$$H(k) = P(k) - P(k) C^T [C P(k) C^T + R]^{-1} C P(k)$$

$$\uparrow$$

$\ell \times \ell$

Kalman Filter Algorithm

1. assume $\hat{x}(0)$, $H(0)$

   select Q, R

2. $P(1) = A(0)\ \mathrm{H}(0)\ A^T(0) + G(0)\ QG^T(0)$

3. $K(1) = P(1)C^T \left[CP(1)C^T + R \right]^{-1}_{\ell x \ell}$

4. $\hat{x}(1) = A(0)\hat{x}(0) + K(1)\left[y(1) - CA\hat{x}(0) \right]$

update

$$H(1) = P(1) - K(1)CP(1)$$

return to 2. to generate $P(2)$, $K(2)$

n.b. If A, G are not functions of k,

$P(k)$, $K(k)$ can be generated ahead of time.

Kalman Filter Extension

Extensions

$$x(k) = A(k-1)\hat{x}(k-1) + B(k-1)u(k-1) + G(k-1)w(k-1)$$

Kalman filter becomes

$$\hat{x}(k) = A(k-1)\hat{x}(k-1) + B(k-1)u(k-1) + \mathrm{K}(k)\left[y(k) - C[A(k-1)\hat{x}(k-1) + B(k-1)u(k-1)]\right]$$

Recursive update for *P*:

$$P(k) = P(k-i) - P(k-1)C^T[CP(k-1)C^T + R]^{-1}CP(k)$$

Properties of Kalman Filter

(a) It provides an unbiased estimate

$$E\left[\tilde{x}(t)\right] = 0 \quad v\ t > o$$

(b) $P(t)$ is the covariance matrix of $\tilde{x}$

$$P = E\left[\tilde{\mathrm{x}}\ \ \tilde{\mathrm{x}}^T\right]$$

$P(t)$ found by soln of Riccati eqn. and does not depend on $y(t)$ – can be calculated a priori

note that $J(t) = \sum_{i=1}^{n} E(\tilde{x}_i^{\,2}(t))$ (variance)

$$= \sum_{i=1}^{n} P_{ii}(t) = tr(P)$$

Properties of Kalman Filter (cont'd)

(c) Kalman filter is a linear, minimum variance estimator

linear o.d.e. relating $\hat{x}$ *to* $y(t)$

For non-white (colored) noise, optimal estimator is not necessarily linear

(d) For long times $(t \to \infty)$

$$K(t) \to \overline{K}$$

$$P(t) \to \overline{P}$$ s.s. Riccati eqn.

don't have to update gain matrix

(e) note similarity to LQP)
(Kalman filter uses initial condition)

Properties of Kalman Filter (cont'd)

(f) Extension of K.F. to nonlinear systems (involves successive linearization of state eqn.)

(g) Q, R. difficult to estimate a priori, but can be used as design parameters
(relative values of Q, R are important)

(h) large $Q \Rightarrow$ large K

(implies process noise is large)

large R (small Q) $\Rightarrow$ small K
(implies measurement noise is large)