Outline:

- why is Kalman Filter famous
- some common skills
- key concepts of Kalman Filter
- multirate Kalman Filter
- about whiteness

1 History of Kalman and his Kalman Filter

- Rudolf Kalman:
  - obtained B.S. in 1953 and M.S. in 1954 from MIT, and Ph.D. in 1957 from Columbia University, all in electrical engineering
  - developed and implemented Kalman Filter (KF) in 1960, during the Apollo program, and furthermore in various famous programs including the NASA Space Shuttle, Navy submarines, etc.
  - Kalman Filter was initially met with vast skepticism; Kalman had to publish the result first in Mechanical Engineering
  - was awarded the National Medal of Science on Oct. 7, 2009 from U.S. president Barack Obama
- Why is Kalman Filter famous: it is a recursive algorithm and suits well for various online implementations

2 Common skills in stochastic control

Assume x is Gaussian distributed.

- If

  $$y = Ax + B$$

  then

  $$X_{yy} = E\left[(y - Ey)(y - Ey)^T\right] = AX_{xx}A^T$$

- If

  $$y = Ax + B$$
  $$y' = A'x + B'$$

  then

  $$X_{yy'} = AX_{xx}\left(A'\right)^T, \; X_{y'y} = A'X_{xx}(A)^T$$

- If

  $$y = Ax + Bv$$
  $$y' = A'x + B'v$$

  and v is Gaussian distributed, then

$$X_{yy'} = AX_{xx}\left(A'\right)^T + AX_{xv}\left(B'\right)^T + BX_{vx}\left(A'\right)^T + BX_{vv}\left(B'\right)^T$$

A, B, A', B' above are all constant matrices.

3 Key concepts of least square estimation

The simplest problem of least square estimation aims at minimizing the cost:

$$J = E\left[\|x-\hat{x}\|_2^2|_{y=y_1}\right] \tag{1}$$

where y is correlated with x.

Solution: (note that $\hat{x}$ is deterministic and that$\|x-\hat{x}\|_2^2$ is quadratic in $\hat{x}$) Setting $\nabla_{\hat{x}}J=0$, we have

$$\hat{x} = E\left[x|y=y_1\right] = m_{x|y=y_1}$$

Properties:

- unbiased estimation:

$$E\left[\hat{x}\right] = E\left[x\right]$$

- if x and y are Gaussian, then
  (1)

$$\boxed{\hat{x} = m_x + X_{xy}X_{yy}^{-1}\left(y_1-m_y\right)} \tag{2}$$

  (2) Estimation error

$$e = x-\hat{x} = x-m_x-X_{xy}X_{yy}^{-1}\left(y_1-m_y\right)$$

$$\boxed{Var\left(e\right) = X_{xx} - \underbrace{X_{xy}X_{yy}^{-1}X_{yx}}_{>0}} \tag{3}$$

  The variance has reduced and it is unbiased estimation. Hence the estimation has been improved, i.e., $E\left[x|y=y_1\right]$ is a better estimate than $E\left[x\right]$.

The case with multiple measurements (closely related to Kalman Filter):

- Concepts of Kalman Filter: (hidden) state estimation from observed/measured outputs

$$\min E\left[\|x\left(k\right)-\hat{x}\left(k\right)\|_2^2|_{y(0),\ldots,y(k)}\right]$$

  which is an extended problem of (1).

- The case where $x\left(k\right)$ is Gaussian is particularly important and will be the focus of the remaining discussions

- When we have two measurements, from the lectures and the last discussion, we know that the best estimate is

$$\begin{aligned} E\left[x|_{y,z}\right] &= E\left[x|_y\right] + E\left[\tilde{x}|_y \middle| \tilde{z}|_y\right] \\ X_{\tilde{x}|_{y,z}\tilde{x}|_{y,z}} &= X_{\tilde{x}|_y\tilde{x}|_y} - X_{\tilde{x}|_y\tilde{z}|_y}X_{\tilde{z}|_y\tilde{z}|_y}^{-1}X_{\tilde{z}|_y\tilde{x}|_y}\end{aligned}$$

  Again, we see that $X_{\tilde{x}|_{y,z}\tilde{x}|_{y,z}} < X_{\tilde{x}|_y\tilde{x}|_y}$, namely, $E\left[x|_{y,z}\right]$ is a better estimation than $E\left[x|_y\right]$.

- When we have multiple measurements in a general linear system*[1]

$$\begin{gathered} x\left(k+1\right) = Ax\left(k\right)+Bu\left(k\right)+B_w w\left(k\right) \\ y\left(k\right) = Cx\left(k\right)+v\left(k\right)\end{gathered}$$

  where

$$\begin{gathered} E\left[w\left(k\right)\right] = 0 \\ E\left[v\left(k\right)\right] = 0 \\ E\left[w\left(k\right)w^T\left(j\right)\right] = W\left(k\right)\delta_{kj} \\ E\left[v\left(k\right)v^T\left(j\right)\right] = V\left(k\right)\delta_{kj} \\ E\left[w\left(k\right)v^T\left(j\right)\right] = 0\ \forall k,j\end{gathered}$$

[1] Extended topic. Please see me for details about the derivation.

We have the best estimate is given by

$$\begin{aligned}
\hat{x}(k)|_{y(0),\dots y(k)} &= E\left[x(k)|_{y(0),\dots y(k)}\right] \\
&= E\left[x(k)|_{y(0),\dots y(k-1)}\right] + E\left[\tilde{x}|_{y(0),\dots,y(k-1)}\middle|\,\tilde{y}(k)|_{y(0),\dots,y(k-1)}\right] \\
&= \hat{x}(k)|_{y(0),\dots,y(k-1)} + \\
& X_{\tilde{x}|_{y(0),\dots,y(k-1)}\tilde{y}(k)|_{y(0),\dots,y(k-1)}} X^{-1}_{\tilde{y}(k)|_{y(0),\dots,y(k-1)}\tilde{y}(k)|_{y(0),\dots,y(k-1)}} \left(\tilde{y}(k)|_{y(0),\dots,y(k-1)} - E\left[\tilde{y}(k)|_{y(0),\dots,y(k-1)}\right]\right)
\end{aligned}$$

Introduce some notations

$$\begin{aligned}
\hat{x}(k|k-1) &= \hat{x}(k)|_{y(0),\dots y(k-1)} \\
\hat{x}(k|k) &= \hat{x}(k)|_{y(0),\dots y(k)} \\
M(k) &= E\left[\tilde{x}(k)|_{y(0),\dots y(k-1)}\tilde{x}^T(k)|_{y(0),\dots y(k-1)}\right] \\
Z(k) &= E\left[\tilde{x}(k)|_{y(0),\dots y(k)}\tilde{x}^T(k)|_{y(0),\dots y(k)}\right]
\end{aligned}$$

and notice that

$$\tilde{y}(k)|_{y(0),\dots y(k-1)} = C\tilde{x}(k)|_{y(0),\dots y(k-1)} + v(k) = C\tilde{x}(k|k-1) + v(k)$$

We have

$$\begin{aligned}
X_{\tilde{x}|_{y(0),\dots,y(k-1)}\tilde{y}(k)|_{y(0),\dots,y(k-1)}} &= M(k)C^T \\
X_{\tilde{y}(k)|_{y(0),\dots,y(k-1)}\tilde{y}(k)|_{y(0),\dots,y(k-1)}} &= CM(k)C^T + V(k) \\
E\left[\tilde{y}(k)|_{y(0),\dots,y(k-1)}\right] &= 0
\end{aligned}$$

and hence

$$\boxed{\hat{x}(k|k) = \hat{x}(k|k-1) + \underbrace{M(k)C^T\left(CM(k)C^T + V(k)\right)^{-1}}_{F(k)}\left(y(k) - C\hat{x}(k|k-1)\right)}$$

The above is the key update equation for Kalman Filter.

- Now for the variance update:

$$\begin{aligned}
& E\left[\tilde{x}(k)|_{y(0),\dots y(k)}\tilde{x}(k)^T|_{y(0),\dots y(k)}\right] \\
=& E\left[\tilde{x}(k)|_{y(0),\dots y(k-1)}\tilde{x}(k)^T|_{y(0),\dots y(k-1)}\right] \\
& - X_{\tilde{x}(k)|_{y(0),\dots y(k-1)}\tilde{y}(k)|_{y(0),\dots y(k-1)}} X^{-1}_{\tilde{y}(k)|_{y(0),\dots y(k-1)}\tilde{y}(k)|_{y(0),\dots y(k-1)}} X_{\tilde{y}(k)|_{y(0),\dots y(k-1)}\tilde{x}(k)|_{y(0),\dots y(k-1)}} \\
=& M(k) - M(k)C^T\left(CM(k)C^T + V(k)\right)^{-1}CM(k)
\end{aligned}$$

namely

$$\boxed{Z(k) = M(k) - M(k)C^T\left(CM(k)C^T + V(k)\right)^{-1}CM(k)}$$

- The connection between $Z(k)$ and $M(k)$:

$$\begin{aligned}
x(k) &= Ax(k-1) + Bu(k-1) + B_w w(k-1) \\
\Rightarrow \hat{x}(k|k-1) &= A\hat{x}(k-1|k-1) + Bu(k-1) \\
\Rightarrow \tilde{x}(k|k-1) &= A\tilde{x}(k-1|k-1) + B_w w(k-1) \\
\Rightarrow M(k) &= AZ(k-1)A^T + B_w W(k) B_w^T,\ M(0) = X_0
\end{aligned}$$

- We have thus obtained a recursive equation for optimally estimating the state $x(k)$ in the scope of least squares. The equations look long but we have actually been using just the basic definitions; (2) and (3); and the skills introduced in Section 2.

- Several important remarks:
  - Kalman Filter is a linear time-varying filter
  - $F(k)$, $M(k)$, and $Z(k)$ can be obtained offline first
  - Although Kalman Filter is linear, it is optimal for Gaussian. Nonlinear estimation can only improve the result if the random process is not Gaussian and we consider at least third-order probability distribution functions
  - Kalman Filter works for time-varying systems

4 Multirate Kalman Filter

Consider the system

$$x(k+1) = Ax(k) + Bu(k) + B_w w(k)$$

where $u(k)$ is the control input, $x(k)$ is the state vector, and $w(k)$ is the Gaussian random input. Additionally, $x(0)$ and $w(k)$ are independent, and

$$E[x(0)] = x_0,\ E\left[(x(0)-x_0)(x(0)-x_0)^T\right] = X_0,\ E[w(k)] = 0,\ E\left[w(k)w^T(j)\right] = W\delta_{kj}$$

Key assumption: the measurement $y(k) = Cx(k) + v(k)$ is taken at every even k, i.e., $k = 0, 2, 4, \ldots$. So the measurement is obtained 2 times slower than the control loop. The Gaussian noise $v(k)$ is independent of $x(0)$ and $w(j)$ for all j and all even k. In addition,

$$E[v(k)] = 0,\ E\left[v(k)v^T(j)\right] = V\delta_{kj},\ \text{for } k = 0, 2, 4, \ldots;\ j = 0, 2, 4, \ldots$$

Find the Kalman filter for this system.

Solution: to let the state estimates updated "at the same rate as the measurement", we need to rewrite the system equations as

$$\begin{aligned}
x(k+2) &= Ax(k+1) + Bu(k+1) + B_w w(k+1) \\
&= A^2 x(k) + [AB, B]\begin{bmatrix} u(k) \\ u(k+1) \end{bmatrix} + [AB_w, B_w]\begin{bmatrix} w(k) \\ w(k+1) \end{bmatrix} \\
&\triangleq \bar{A}x(k) + \bar{B}\bar{u}(k) + \bar{B}_w \bar{w}(k) \\
y(k+2) &= Cx(k+2) + v(k+2)
\end{aligned}$$

then we have

$$E[\bar{w}(k)] = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ E\left[\bar{w}(k)\bar{w}^T(j)\right] = \begin{bmatrix} W & 0 \\ 0 & W \end{bmatrix}\delta_{kj},\ \text{for all even } k\text{and all even } j$$

From equations (KF19-23), we then have

$$\begin{aligned}
\hat{x}(k+2|k+2) &= \hat{x}(k+2|k) + F(k+2)\left[y(k+2) - C\hat{x}(k+2|k)\right] \\
\hat{x}(k+2|k) &= \bar{A}\hat{x}(k|k) + \bar{B}\bar{u}(k),\ \hat{x}(0|-2) = x_0 \\
F(k+2) &= M(k+2)C^T\left[CM(k+2)C^T + V\right]^{-1} \\
M(k+2) &= \bar{A}Z(K)\bar{A}^T + \bar{B}_w\begin{bmatrix} W & 0 \\ 0 & W \end{bmatrix}\bar{B}_w^T,\ M(0) = X_0 \\
Z(k+2) &= M(k+2) - M(k+2)C^T\left[CM(K+2)C^T + V\right]^{-1}CM(k+2)
\end{aligned}$$

5 Colored input noise

Consider the system $G(z) = \frac{1}{z-1}$, with the input $u(k)$, input disturbance $\bar{w}(k)$, output $y(k)$, and output disturbance $v(k)$. All the standard conditions for Kalman Filter hold except that the input noise $\bar{w}(k)$ is not white but a zero mean WSS colored noise with power spectral density $\Phi_{\bar{w}\bar{w}}(\omega) = \frac{5+4\cos\omega}{5-4\cos\omega}$.

To apply the standard Kalman Filter, we first model $\bar{w}(k)$ as the output of a stable linear system with a zero mean WSS white noise as the input. To do so, we can factorize $\Phi_{\bar{w}\bar{w}}(\omega)$ as

$$\Phi_{\bar{w}\bar{w}}(\omega) = \frac{5 + 2e^{j\omega} + 2e^{-j\omega}}{5 - 2e^{j\omega} - 2e^{-j\omega}} = \frac{\left(2e^{j\omega}+1\right)\left(2e^{-j\omega}+1\right)}{\left(2e^{j\omega}-1\right)\left(2e^{-j\omega}-1\right)} = G\left(e^{j\omega}\right)G\left(e^{-j\omega}\right)$$

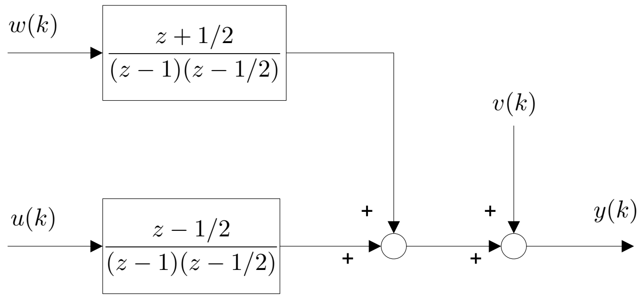
where $G(z) = \frac{2z+1}{2z-1}$ and $G\left(e^{j\omega}\right) = G(z)|_{z=e^{j\omega}}$.

Then $\bar{w}(k)$ can be considered as the output of the system

$$w(k) \longrightarrow \boxed{\frac{2z+1}{2z-1}} \longrightarrow \bar{w}(k)$$

where $w(k)$ is a WSS white noise with zero mean and unit variance.

The system block diagram becomes



One state-space realization of the above system is

$$\begin{aligned} x(k+1) &= \begin{bmatrix} 3/2 & 1 \\ -1/2 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ -1/2 \end{bmatrix} u(k) + \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} w(k) \\ y(k) &= \begin{bmatrix} 1 & 0 \end{bmatrix} x(k) + v(k) \end{aligned}$$

Now all the standard conditions hold for this augmented system and we can apply the Kalman Filter equations in the reader.