Stochastic State Estimation (Kalman Filter)

Big picture

why are we learning this?

- state estimation in deterministic case:

$$\text{Plant: } x(k+1) = Ax(k) + Bu(k), \; y(k) = Cx(k)$$
$$\text{Observer: } \hat{x}(k+1) = A\hat{x}(k) + Bu(k) + L(y(k) - C\hat{x}(k))$$

- L designed based on the error $(e(k) = x(k) - \hat{x}(k))$ dynamics:

$$e(k+1) = (A - LC)\, e(k) \tag{1}$$

  to reach fast convergence of $\lim_{k\to\infty} e(k) = 0$

- L is not optimal when there is noise in the plant; actually $\lim_{k\to\infty} e(k) = 0$ isn't even a valid goal when there is noise
- Kalman Filter provides optimal state estimation under input and output noises

Problem statement

plant:
$$x(k+1) = A(k)\,x(k) + B(k)\,u(k) + B_w(k)\,w(k)$$
$$y(k) = C(k)\,x(k) + v(k)$$

- $w(k)$–s-dimensional input noise; $v(k)$–r-dimensional measurement noise; $x(0)$–unknown initial state
- assumptions: $x(0)$, $w(k)$, and $v(k)$ are independent and Gaussian distributed; $w(k)$ and $v(k)$ are white:

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

Problem statement

- goal:

$$\text{minimize } \mathrm{E}\left[||x(k) - \hat{x}(k)||^2 \Big|_{Y_j} \right], \; Y_j = \{y(0), y(1), \ldots, y(j)\}$$

- solution:

$$\hat{x}(k) = \mathrm{E}\left[x(k) \,|\, Y_j\right]$$

- three classes of problems:
  - $k > j$: prediction problem
  - $k = j$: filtering problem
  - $k < j$: smoothing problem

History

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 in 1960, during the Apollo program, and furthermore in various famous programs including the NASA Space Shuttle, Navy submarines, etc.
- was awarded the National Medal of Science on Oct. 7, 2009 from U.S. president Barack Obama

Useful facts

assume x is Gaussian distributed

- if $y = Ax + B$ then

$$\begin{cases} X_{xy} = \mathrm{E}\left[(x - \mathrm{E}[x])(y - \mathrm{E}[y])^T\right] & = X_{xx}A^T \\ X_{yy} = \mathrm{E}\left[(y - \mathrm{E}[y])(y - \mathrm{E}[y])^T\right] & = AX_{xx}A^T \end{cases} \tag{2}$$

- if $y = Ax + B$ and $y^{'} = A^{'}x + B^{'}$ then

$$X_{yy^{'}} = AX_{xx}\left(A^{'}\right)^T,\ X_{y^{'}y} = A^{'}X_{xx}A^T \tag{3}$$

- if $y = Ax + Bv$; v is Gaussian and independent of x, then

$$X_{yy} = AX_{xx}A^T + BX_{vv}B^T \tag{4}$$

- if $y = Ax + Bv$, $y^{'} = A^{'}x + B^{'}v$; v is Gaussian and dependent of x, then

$$X_{yy^{'}} = AX_{xx}\left(A^{'}\right)^T + AX_{xv}\left(B^{'}\right)^T + BX_{vx}\left(A^{'}\right)^T + BX_{vv}\left(B^{'}\right)^T \tag{5}$$

Derivation of Kalman Filter

- goal:

$$\text{minimize } \mathrm{E}\left[||x(k) - \hat{x}(k)||^2 \big|_{Y_k} \right], \; Y_k = \{y(0), y(1), \dots, y(k)\}$$

- the best estimate is the conditional expectation

$$\begin{aligned} \mathrm{E}[x(k)|Y_k] &= \mathrm{E}[x(k)|\{Y_{k-1}, y(k)\}] \\ &= \mathrm{E}[x(k)|Y_{k-1}] + \mathrm{E}\left[\tilde{x}(k)|_{Y_{k-1}} \big| \tilde{y}(k)|_{Y_{k-1}} \right] \end{aligned}$$

- introduce some notations:

$$\begin{aligned} \text{a priori estimation } \hat{x}(k|k-1) &= \mathrm{E}[x(k)|Y_{k-1}] = \hat{x}(k)|_{y(0),\dots y(k-1)} \\ \text{a posteriori estimation } \hat{x}(k|k) &= \mathrm{E}[x(k)|Y_k] = \hat{x}(k)|_{y(0),\dots y(k)} \\ \text{a priori covariance } M(k) &= \mathrm{E}\left[\tilde{x}(k)|_{Y_{k-1}} \tilde{x}^T(k)|_{Y_{k-1}} \right] \\ \text{a posteriori covariance } Z(k) &= \mathrm{E}\left[\tilde{x}(k)|_{Y_k} \tilde{x}^T(k)|_{Y_k} \right] \end{aligned}$$

Derivation of Kalman Filter

KF gain update

to get $\mathrm{E}\left[\tilde{x}(k)|_{Y_{k-1}}\middle|\tilde{y}(k)|_{Y_{k-1}}\right]$ in

$$\mathrm{E}\left[x(k)|Y_k\right] = \mathrm{E}\left[x(k)|Y_{k-1}\right] + \mathrm{E}\left[\tilde{x}(k)|_{Y_{k-1}}\middle|\tilde{y}(k)|_{Y_{k-1}}\right]$$

we need $X_{\tilde{x}(k)|_{Y_{k-1}}\tilde{y}(k)|_{Y_{k-1}}}$ and $X^{-1}_{\tilde{y}(k)|_{Y_{k-1}}\tilde{y}(k)|_{Y_{k-1}}}$

$y(k) = C(k)x(k) + v(k)$ gives

$$\begin{aligned} \hat{y}(k)|_{Y_{k-1}} &= C(k)\hat{x}(k|k-1) + \hat{v}(k)|_{Y_{k-1}} = C(k)\hat{x}(k|k-1) \\ \Rightarrow \tilde{y}(k)|_{Y_{k-1}} &= C(k)\tilde{x}(k|k-1) + v(k) \end{aligned}$$

hence

$$X_{\tilde{x}(k)|_{Y_{k-1}}\tilde{y}(k)|_{Y_{k-1}}} = M(k)C^T(k) \tag{6}$$

$$X_{\tilde{y}(k)|_{Y_{k-1}}\tilde{y}(k)|_{Y_{k-1}}} = C(k)M(k)C^T(k) + V(k) \tag{7}$$

Derivation of Kalman Filter

KF gain update

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

unbiased estimation: $\mathrm{E}\left[\hat{x}(k|k-1)\right] = \mathrm{E}[x] \Rightarrow$

$$\mathrm{E}\left[\tilde{y}(k)|_{Y_{k-1}}\right] = \mathrm{E}\left[\tilde{x}(k)|_{Y_{k-1}}\right] + \mathrm{E}\left[v(k)|_{Y_{k-1}}\right] = 0$$

thus

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

Derivation of Kalman Filter

KF gain update

$$\mathrm{E}\left[x(k)\,|\,Y_k\right] = \mathrm{E}\left[x(k)\,|\,Y_{k-1}\right] + \mathrm{E}\left[\tilde{x}(k)\,|_{Y_{k-1}}\middle|\,\tilde{y}(k)\,|_{Y_{k-1}}\right]$$

now becomes

$$\begin{aligned}\hat{x}(k|k) &= \hat{x}(k|k-1) \\ &\quad + \underbrace{M(k)\,C^T\left(CM(k)\,C^T + V(k)\right)^{-1}}_{F(k)}\left(y(k) - C\hat{x}(k|k-1)\right)\end{aligned}$$

namely

$$\boxed{\begin{cases}\hat{x}(k|k) & = \hat{x}(k|k-1) + F(k)\left(y(k) - C(k)\,\hat{x}(k|k-1)\right) \\ F(k) & = M(k)\,C^T(k)\left(C(k)\,M(k)\,C^T(k) + V(k)\right)^{-1}\end{cases}} \qquad (8)$$

Derivation of Kalman Filter

KF covariance update

now for the variance update:

$$\mathrm{E}\left[\tilde{x}(k)|_{Y_k}\tilde{x}(k)^T|_{Y_k}\right] = \mathrm{E}\left[\tilde{x}(k)|_{\{Y_{k-1},y(k)\}}\tilde{x}(k)^T|_{\{Y_{k-1},y(k)\}}\right]$$

$$= \mathrm{E}\left[\tilde{x}(k)|_{Y_{k-1}}\tilde{x}(k)^T|_{Y_{k-1}}\right]$$

$$- X_{\tilde{x}(k)|_{Y_{k-1}}\tilde{y}(k)|_{Y_{k-1}}} X^{-1}_{\tilde{y}(k)|_{Y_{k-1}}\tilde{y}(k)|_{Y_{k-1}}} X_{\tilde{y}(k)|_{Y_{k-1}}\tilde{x}(k)|_{Y_{k-1}}}$$

or, using the introduced notations,

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

Derivation of Kalman Filter

KF covariance update

the connection between $Z(k)$ and $M(k)$:

$$\begin{aligned} x(k) &= A(k-1)\,x(k-1) + B(k-1)\,u(k-1) + B_w(k-1)\,w(k-1) \\ \Rightarrow \hat{x}(k|k-1) &= A(k-1)\,\hat{x}(k-1|k-1) + B(k-1)\,u(k-1) \\ \Rightarrow \tilde{x}(k|k-1) &= A(k-1)\,\tilde{x}(k-1|k-1) + B_w(k-1)\,w(k-1) \end{aligned}$$

thus $M(k) = \mathrm{Cov}\left[\tilde{x}(k|k-1)\right]$ is [using uesful fact (4)]

$$\boxed{M(k) = A(k-1)\,Z(k-1)\,A^T(k-1) + B_w(k-1)\,W(k-1)\,B_w^T(k-1)}$$

with $M(0) = \mathrm{E}\left[\tilde{x}(0|-1)\,\tilde{x}(0|-1)^T\right] = X_0$

The full set of KF equations

$$
\begin{aligned}
\hat{x}(k|k) &= \hat{x}(k|k-1) + F(k)\overbrace{\left[y(k) - C(k)\,\hat{x}(k|k-1)\right]}^{e_y(k)} \\
\hat{x}(k|k-1) &= A(k-1)\,\hat{x}(k-1|k-1) + B(k-1)\,u(k-1) \\
F(k) &= M(k)C^T(k)\left[C(k)\,M(k)C^T(k) + V(k)\right]^{-1} \\
M(k) &= A(k-1)\,Z(k-1)A^T(k-1) + B_w(k-1)\,W(k-1)\,B_w^T(k-1) \\
Z(k) &= M(k) - M(k)\,C^T(k)\ldots \\
&\qquad \cdot\left(C(k)\,M(k)\,C^T(k) + V(k)\right)^{-1} C(k)\,M(k) \\
&= (I - F(k)\,C(k))\,M(k)
\end{aligned}
$$

with initial conditions $\hat{x}(0|-1) = x_o$ and $M(0) = X_0$.

The full set of KF equations

in a shifted index:

$$
\begin{aligned}
\hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + F(k+1)\left[y(k+1) - C\,(k+1)\,\hat{x}(k+1|k)\right] \\
\hat{x}(k+1|k) &= A\,(k)\,\hat{x}(k|k) + B\,(k)\,u(k) \\
F(k+1) &= M(k+1)C^{T}\,(k+1)\left[C\,(k+1)\,M(k+1)C^{T}\,(k+1) + V\,(k+1)\right]^{-1} \\
M(k+1) &= A\,(k)\,Z(k)A^{T}\,(k) + B_{w}\,(k)\,W\,(k)\,B_{w}^{T}\,(k) && (9) \\
Z(k+1) &= M\,(k+1) - M\,(k+1)\,C^{T}\,(k+1)\ldots && (10) \\
&\quad \cdot\left(C\,(k+1)\,M\,(k+1)\,C^{T}\,(k+1) + V\,(k+1)\right)^{-1} C\,(k+1)\,M\,(k+1) \\
&= \left(I - F\,(k+1)\,C\,(k+1)\right) M\,(k+1) && (11)
\end{aligned}
$$

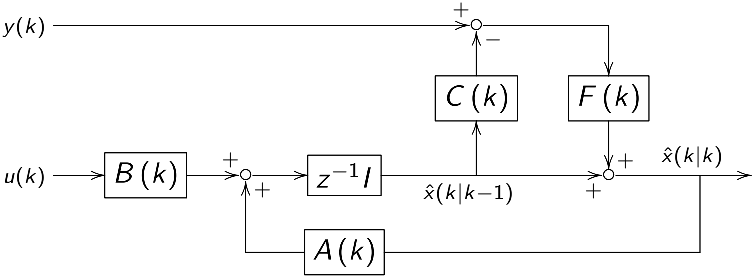
combining (9) and (10) gives the Riccati equation:

$$
\begin{aligned}
M\,(k+1) &= A\,(k)\,M\,(k)\,A^{T}\,(k) + B_{w}\,(k)\,W\,(k)\,B_{w}^{T}\,(k) \\
&- A\,(k)\,M\,(k)\,C^{T}\,(k)\left[C\,(k)\,M\,(k)\,C^{T}\,(k) + V\,(k)\right]^{-1} C\,(k)\,M\,(k)\,A^{T}\,(k) && (12)
\end{aligned}
$$

The full set of KF equations

Several remarks

- $F(k)$, $M(k)$, and $Z(k)$ can be obtained offline first
- Kalman Filter (KF) is linear, and optimal for Gaussian. More advanced nonlinear estimation won't improve the results here.
- KF works for time-varying systems
- the block diagram of KF is:



Steady-state KF

assumptions:

- system is time-invariant: $A(k)$, $B(k)$, $B_w(k)$, and $C(k)$ are constant;
- noise is stationary: $V \succ 0$ and $W \succ 0$ do not depend on time.

KF equations become:

$$
\begin{aligned}
\hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + F(k+1)\left[y(k+1) - C\hat{x}(k+1|k)\right] \\
&= A\hat{x}(k|k) + Bu(k) + F(k+1)\left[y(k+1) - C\hat{x}(k+1|k)\right] \\
F(k+1) &= M(k+1)C^T\left[CM(k+1)C^T + V\right]^{-1} \\
M(k+1) &= AZ(k)A^T + B_w W B_w^T; \quad M(0) = X_0 \\
Z(k+1) &= M(k+1) - M(k+1)C^T\left[CM(k+1)C^T + V\right]^{-1} CM(k+1)
\end{aligned}
$$

with Riccati equation (RE):

$$
M(k+1) = AM(k)A^T + B_w W B_w^T - AM(k)C^T\left[CM(k)C^T + V\right]^{-1} CM(k)A^T
$$

Steady-state KF

if

- (A, C) is observable or detectable
- (A, B_w) is controllable (disturbable) or stabilizable

then $M(k)$ in the RE converges to some steady-state value M_s and KF can be implemented by

$$\begin{aligned}
\hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + F_s \left[y(k+1) - C\hat{x}(k+1|k) \right] \\
\hat{x}(k+1|k) &= A\hat{x}(k|k) + Bu(k) \\
F_s &= M_s C^T \left[C M_s C^T + V \right]^{-1}
\end{aligned}$$

M_s is the positive definite solution of the algebraic Riccati equation:

$$M_s = A M_s A^T + B_w W B_w^T - A M_s C^T \left[C M_s C^T + V \right]^{-1} C M_s A^T$$

Duality with LQ

The steady-state condition is obtained by comparing the RE in LQ and KF

discrete-time LQ:

$$P(k) = A^T P(k+1)A - A^T P(k+1)B[R + B^T P(k+1)B]^{-1} B^T P(k+1)A + Q$$

discrete-time KF (12):

$$M(k+1) = AM(k)A^T - AM(k)C^T \left[CM(k)C^T + V\right]^{-1} CM(k)A^T + B_w W B_w^T$$

discrete-time LQ	discrete-time KF
A	A^T
B	C^T
C	B_w
R	V
$Q = C^T C$	$B_w W B_w^T$
P	M
backward recursion	forward recursion

Duality with LQ

discrete-time LQ	discrete-time KF
A	A^T
B	C^T
C	B_w
$Q = C^T C$	$B_w W B_w^T$

steady-state conditions for discrete-time LQ:

- (A, B) controllable or stabilizable
- (A, C) observable or detectable

steady-state conditions for discrete-time KF:

- $\left(A^T, C^T\right)$ controllable or stabilizable$\Leftrightarrow(A, C)$ observable or detectable
- $\left(A^T, B_w^T\right)$ observable or detectable$\Leftrightarrow(A, B_w)$ controllable or stabilizable

Duality with LQ

discrete-time LQ	discrete-time KF
A	A^T
B	C^T
C	B_w
R	V
$Q = C^T C$	$B_w W B_w^T$
P	M
backward recursion	forward recursion

- LQ: stable closed-loop "A" matrix is

$$A - BK_s = \underline{A - B[R + B^T P_s B]^{-1} B^T P_s A}$$

- KF: stable KF "A" matrix is

$$\begin{aligned}
\hat{x}(k+1|k) &= A\hat{x}(k|k) + Bu(k) \\
&= A\hat{x}(k|k-1) + AF_s\left[y(k) - C\hat{x}(x|k-1)\right] + Bu(k) \\
&= \underline{\left[A - AM_s C^T \left(CM_s C^T + V\right)^{-1} C\right]}\hat{x}(k|k-1) + \ldots
\end{aligned}$$

Purpose of each condition

- (A, C) observable or detectable: assures the existence of the steady-state Riccati solution
- (A, B_w) controllable or stabilizable: assures that the steady-state solution is positive definite and that the KF dynamics is stable

Remark

- KF: stable KF "A" matrix is

$$\begin{aligned}\hat{x}(k+1|k) &= \left[A - AM_sC^T\left(CM_sC^T+V\right)^{-1}C\right]\hat{x}(k|k-1)+\dots\\ &= \underline{(A-AF_sC)}\hat{x}(k|k-1)+\dots\end{aligned}$$

in the form of $\hat{x}(k|k)$ dynamics:

$$\begin{aligned}\hat{x}(k+1|k+1) &= \hat{x}(k+1|k)+F_s\left[y(k+1)-C\hat{x}(k+1|k)\right]\\ &= \underline{(A-F_sCA)}\hat{x}(k|k)+(I-F_sC)\,Bu(k)+F_sy(k+1)\\ &= \left[A - M_sC^T\left(CM_sC^T+V\right)^{-1}CA\right]\hat{x}(k|k)+\dots\end{aligned}$$

- can show that $eig\,(A-AF_sC) = eig\,(A-F_sCA)$

hint: $\det(I+MN) = \det(I+NM) \Rightarrow \det\left[I - z^{-1}A(I-F_sC)\right] = \det\left[I-(I-F_sC)\,z^{-1}A\right]$

Remark

intuition of guaranteed KF stability: ARE $\Rightarrow$ Lyapunov equation

$$
\begin{aligned}
M_s &= AM_sA^T + B_wWB_w^T - AM_sC^T\left[CM_sC^T+V\right]^{-1}CM_sA^T \\
&= AM_sA^T + B_wWB_w^T - A\underbrace{M_sC^T\left[CM_sC^T+V\right]^{-1}}_{F_s}\left[CM_sC^T+V\right]\underbrace{\left[CM_sC^T+V\right]^{-1}CM_s}_{F_s^T}A^T \\
&= \left(A - AF_sC\right)M_s\left(A-AF_sC\right)^T + 2AF_sCM_sA^T - AF_sCM_sC^TF_s^TA^T \\
&\qquad + B_wWB_w^T - AF_s\left[CM_sC^T+V\right]F_s^TA^T \\
&= \left(A - AF_sC\right)M_s\left(A-AF_sC\right)^T + AF_sVF_s^TA^T + B_wWB_w^T
\end{aligned}
$$

$$
\Longleftrightarrow \left(A - AF_sC\right)M_s\left(A-AF_sC\right)^T - M_s = -AF_sVF_s^TA^T - B_wWB_w^T
$$

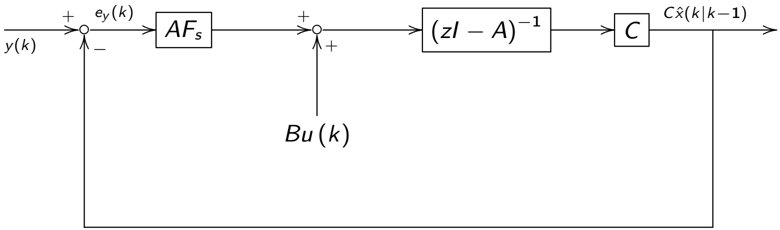
which is a Lyapunov equation with the right hand side being negative semidefinite and $M_s \succ 0$.

Return difference equation

KF dynamics

$$
\begin{aligned}
\hat{x}(k+1|k) &= A\hat{x}(k|k) + Bu(k) \\
&= A\hat{x}(k|k-1) + AF_s\left[y\left(k\right) - C\hat{x}\left(x|k-1\right)\right] + Bu\left(k\right)
\end{aligned}
$$

$$
\left[zI - A\right]\mathcal{Z}\left\{\hat{x}\left(k|k-1\right)\right\} = AF_s\mathcal{Z}\left\{y(k) - C\hat{x}(k|k-1)\right\} + B\mathcal{Z}\left\{u(k)\right\}
$$



let $G(z) = C(zI - A)^{-1}B_w$
ARE $\Rightarrow$ return difference equation (RDE) (see course reader)

$$
\left[I + C(zI - A)^{-1}AF_s\right]\left(V + CM_sC^T\right)\left[I + C(z^{-1}I - A)^{-1}AF_s\right]^T = V + G(z)WG^T(z^{-1})
$$

Symmetric root locus for KF

- KF eigenvalues:

$$\det\left[I + C\underline{(zI-A)^{-1}AF_s}\right] = \det\left[I + \underline{(zI-A)^{-1}AF_s}C\right]$$
$$= \frac{\det\left(zI - A + AF_sC\right)}{\det\left(zI-A\right)} \triangleq \frac{\beta\left(z\right)}{\phi\left(z\right)}$$

- taking determinants in RDE gives

$$\beta(z)\beta(z^{-1}) = \phi(z)\phi(z^{-1})\frac{\det\left(V + G(z)WG^T(z^{-1})\right)}{\det\left(V + CMC^T\right)}$$

- single-output case: KF poles come from $\beta(z)\beta(z^{-1}) = 0$, i.e.

$$\det\left(V + G(z)WG^T(z^{-1})\right) = V\left(1 + G(z)\frac{W}{V}G^T(z^{-1})\right) = 0$$

- $W/V \to 0$: KF poles $\to$ stable poles of $G\left(z\right)G^T\left(z^{-1}\right)$
- $W/V \to \infty$: KF poles $\to$ stable zeros of $G\left(z\right)G^T\left(z^{-1}\right)$

Continuous-time KF

summary of solutions

system:

$$\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t)$$
$$y(t) = Cx(t) + v(t)$$

assumptions: same as discrete-time KF

aim: minimize $J = \left. ||x(t) - \hat{x}(t)||_2^2 \right|_{\{y(\tau):0\leq\tau\leq t\}}$

continuous-time KF:

$$\frac{\mathrm{d}\hat{x}(t|t)}{\mathrm{d}t} = A\hat{x}(t|t) + Bu(t) + F(t)\left[y(t) - C\hat{x}(t|t)\right], \ \hat{x}(0|0) = x_0$$

$$F(t) = M(t)\,C^T V^{-1}$$

$$\frac{\mathrm{d}M(t)}{\mathrm{d}t} = AM(t) + M(t)\,A^T + B_w W B_w^T - M(t)\,C^T V^{-1} CM(t), \ M(0) = X_0$$

Continuous-time KF: steady state

assumptions: (A, C) observable or detectable;
(A, B_w) controllable or stabilizable

asymptotically stable steady-state KF:

$$\frac{\mathrm{d}\hat{x}(t|t)}{\mathrm{d}t} = A\hat{x}(t|t) + Bu(t) + F_s[y(t) - C\hat{x}(t|t)]$$

$$F_s = M_s C^T V^{-1}$$

$$AM_s + M_s A^T + B_w W B_w^T - M_s C^T V^{-1} C M_s = 0$$

duality with LQ:

Continuous-Time LQ
$A^T P_s + P_s A + Q - P_s B R^{-1} B^T P_s = 0$ $K = R^{-1} B^T P_s$

Continuous-time KF: return difference equality

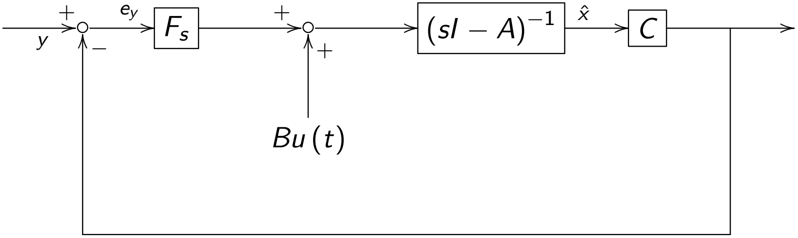
analogy to LQ gives the return difference equality:

$$\left[I + C\left(sI - A\right)^{-1} F_s\right] V \left[I + F_s^T \left(-sI - A\right)^{-T} C^T\right] = V + G\left(s\right) W G^T\left(-s\right)$$

where $G\left(s\right) = C\left(sI - A\right)^{-1} B_w$, hence:

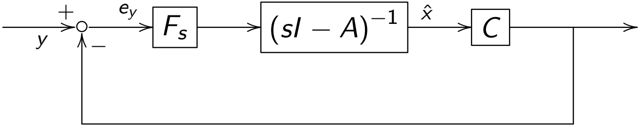
$$\left[I + C\left(j\omega I - A\right)^{-1} F_s\right] V \left[I + C\left(-j\omega I - A\right)^{-1} F_s\right]^T = V + G\left(j\omega\right) W G^T\left(-j\omega\right)$$

observation 1: $\frac{\mathrm{d}\hat{x}(t|t)}{\mathrm{d}t} = A\hat{x}\left(t|t\right) + Bu\left(t\right) + F_s \underbrace{\left[y\left(t\right) - C\hat{x}\left(t|t\right)\right]}_{e_y(t)}$



Continuous-time KF: properties

observation 1:



- transfer function from y to e_y: $\left[I + C\left(j\omega I - A\right)^{-1} F_s\right]^{-1}$
- spectral density relation:

$$\Phi_{e_y e_y}(\omega) = \underbrace{\left[I + C\left(j\omega I - A\right)^{-1} F_s\right]^{-1}}_{G_{y \to e_y}(j\omega)} \Phi_{yy}(\omega) \underbrace{\left\{\left[I + C\left(-j\omega I - A\right)^{-1} F_s\right]^{-1}\right\}^T}_{G_{y \to e_y}(-j\omega)^T}$$

Continuous-time KF: properties

observation 2:

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) + B_w w(t) \\ y(t) &= Cx(t) + v(t) \end{cases} \Rightarrow \Phi_{yy}(\omega) = G(j\omega) W G^T(-j\omega) + V$$

from observations 1 and 2:

$$\left[I + C(j\omega I - A)^{-1} F_s\right] V \left[I + C(-j\omega I - A)^{-1} F_s\right]^T = V + G(j\omega) W G^T(-j\omega)$$

thus says

$$\begin{aligned} \Phi_{e_y e_y}(\omega) &= \left[I + C(j\omega I - A)^{-1} F_s\right]^{-1} \Phi_{yy}(\omega) \left\{\left[I + C(-j\omega I - A)^{-1} F_s\right]^{-1}\right\}^T \\ &= V \end{aligned}$$

namely, the estimation error is white!

Continuous-time KF: symmetric root locus

taking determinants of RDE gives:

$$\det\left[I + C\left(sI - A\right)^{-1} F_s\right] \det V \det\left[I + C\left(-sI - A\right)^{-1} F_s\right]^T$$
$$= \det\left[V + G\left(s\right) W G^T\left(-s\right)\right]$$

for single-output systems:

$$\det\left[I + C\left(sI - A\right)^{-1} F_s\right] \det\left[I + C\left(-sI - A\right)^{-1} F_s\right]^T = 1 + G\left(s\right) \frac{W}{V} G^T\left(-s\right)$$

Continuous-time KF: symmetric root locus

the left hand side of

$$\det\left[I + C(sI - A)^{-1} F_s\right] \det\left[I + C(-sI - A)^{-1} F_s\right]^T = 1 + G(s)\frac{W}{V}G^T(-s)$$

determines the KF eigenvalues:

$$\begin{aligned}
\det\left[I + C(sI - A)^{-1} F_s\right] &= \det\left[I + (sI - A)^{-1} F_s C\right] \\
&= \det\left[(sI - A)^{-1}\right] \det\left[sI - A + F_s C\right] \\
&= \frac{\det\left[sI - (A - F_s C)\right]}{\det(sI - A)}
\end{aligned}$$

hence looking at $1 + G(s)\frac{W}{V}G^T(-s)$, we have:

- $W/V \to 0$: KF poles $\to$ stable poles of $G(s)G^T(-s)$
- $W/V \to \infty$: KF poles $\to$ stable zeros of $G(s)G^T(-s)$

Summary

1. Big picture
2. Problem statement
3. Discrete-time KF
   - Gain update
   - Covariance update
   - Steady-state KF
   - Duality with LQ
4. Continuous-time KF
   - Solution
   - Steady-state solution and conditions
   - Properties: return difference equality, symmetric root locus...

*KF equations using other notation systems

- the field of information fusion often uses a different set of notations
- system model:

$$x(k+1) = A(k)\,x(k) + B(k)\,u(k) + w(k)$$
$$z(k) = H(k)\,x(k) + v(k)$$

where $\mathrm{E}\left[w(k)\,w(i)^T\right] = Q(k)\,\delta(k-i)$,
$\mathrm{E}\left[v(k)\,v(i)^T\right] = R(k)\,\delta(k-i)$, and $\mathrm{E}\left[w(k)\,v(i)^T\right] = 0\ \forall k, i$.

- state prediction covariance:

$$P(k+1|k) \triangleq \mathrm{E}\left\{\left(x(k+1) - \hat{x}(k+1|k)\right)\left(x(k+1) - \hat{x}(k+1|k)\right)^T\right\}$$

- Kalman Filter gain: denoted as $K(k)$ (sometimes also as $W(k)$)

*KF equations using other notation systems

- prediction

$$\hat{x}(k|k-1) = A(k-1)\,\hat{x}(k-1|k-1) + B(k-1)\,u(k-1)$$
$$P(k|k-1) = A(k-1)\,P(k-1|k-1)A^T(k-1) + Q(k-1)$$

- correction/update

$$\begin{aligned}
\hat{x}(k|k) &= \hat{x}(k|k-1) + K(k)\left[z(k) - H(k)\,\hat{x}(k|k-1)\right] \\
K(k) &= P(k|k-1)H^T(k)\left[H(k)\,P(k|k-1)H^T(k) + R(k)\right]^{-1} \\
P(k|k) &= P(k|k-1) - P(k|k-1)\,H^T(k)\ldots \\
&\qquad \cdot \left(H(k)\,P(k|k-1)\,H^T(k) + R(k)\right)^{-1} H(k)\,P(k|k-1) \\
&= (I - K(k)\,H(k))\,P(k|k-1)
\end{aligned}$$

with initial conditions $\hat{x}(0|-1) = x_o$ and $P(0|0) = X_0$.

*KF equations using other notation systems

in a shifted index:

$$
\begin{aligned}
\hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + K(k+1)\left[z(k+1) - H(k+1)\,\hat{x}(k+1|k)\right] \\
\hat{x}(k+1|k) &= A(k)\,\hat{x}(k|k) + B(k)\,u(k) \\
K(k+1) &= P(k+1|k)H^T(k+1)\left[H(k+1)\,P(k+1|k)H^T(k+1) + R(k+1)\right]^{-1} \\
P(k+1|k) &= A(k)\,P(k|k)A^T(k) + Q(k) \\
P(k+1|k+1) &= P(k+1|k) - P(k+1|k)\,H^T(k+1)\ldots \\
&\quad \cdot \left(H(k+1)\,P(k+1|k)\,H^T(k+1) + R(k+1)\right)^{-1}\ldots \\
&\quad \cdot H(k+1)\,P(k+1|k) \\
&= \left(I - K(k+1)\,H(k+1)\right)P(k+1|k)
\end{aligned}
$$

*The information form of the KF, aka, information filter

- system model (leveraging the benefits of the two notation systems):

$$x(k+1) = A(k)\,x(k) + B(k)\,u(k) + w(k)$$
$$y(k) = C(k)\,x(k) + v(k)$$

where $\mathrm{E}\left[w(k)\,w(i)^T\right] = Q(k)\,\delta(k-i)$,
$\mathrm{E}\left[v(k)\,v(i)^T\right] = R(k)\,\delta(k-i)$, and $\mathrm{E}\left[w(k)\,v(i)^T\right] = 0\ \forall k, i$.

- state prediction covariance:

$$P(k+1|k) \triangleq \mathrm{E}\left\{(x(k+1) - \hat{x}(k+1|k))\,(x(k+1) - \hat{x}(k+1|k))^T\right\}$$

*The information form of the KF, aka, information filter

- KF:

$$\begin{aligned}
\hat{x}(k|k-1) &= A(k-1)\,\hat{x}(k-1|k-1) + B(k-1)\,u(k-1) \\
P(k|k-1) &= A(k-1)\,P(k-1|k-1)A^T(k-1) + Q(k-1) & \uparrow \text{predict} \\
\hline
\hat{x}(k|k) &= \hat{x}(k|k-1) + K(k)\left[y(k) - C(k)\,\hat{x}(k|k-1)\right] & \downarrow \text{correct} \\
K(k) &= P(k|k-1)C^T(k)\left[C(k)\,P(k|k-1)C^T(k) + R(k)\right]^{-1} \\
P(k|k) &= P(k|k-1) - P(k|k-1)\,C^T(k)\ldots \\
&\qquad \cdot\left(C(k)\,P(k|k-1)\,C^T(k) + R(k)\right)^{-1}\ldots \\
&\qquad \cdot\, C(k)\,P(k|k-1) \\
&= \left(I - K(k)\,C(k)\right)P(k|k-1)
\end{aligned}$$

with initial conditions $\hat{x}(0|-1) = x_o$ and $P(0|0) = X_0$.

*The information form of the KF, aka, information filter

- standard KF form computes the inverse of the innovation covariance matrix, i.e.,

$$\left[C(k)\,P(k|k-1)C^T(k)+R(k)\right]^{-1}$$

  which can be expensive, e.g.:
  - for high-dimension systems and when nondiagonal elements appear
  - when many sensors are available to create a fusion of information
- from the Matrix Inversion Lemma, $\left[C(k)\,P(k|k-1)C^T(k)+R(k)\right]^{-1}$ will involve the inverse of the covariance $P(k|k-1)^{-1}$, the information matrix, aka the Fisher information

*The information form of the KF, aka, information filter

- let's focus on the Fisher information. By using the Matrix Inversion Lemma, we can rewrite the innovation covariance equation as

$$P(k|k)^{-1} = P(k|k-1)^{-1} + C(k)^T R(k)^{-1} C(k)$$

b/c $\left(A + BCB^T\right)^{-1} = A^{-1} - A^{-1}B\left(B^T A^{-1} B + C^{-1}\right)^{-1} B^T A^{-1}$ and hence

$$\left(P_{k|k-1}^{-1} + C_k^T R_k^{-1} C_k\right)^{-1} = P_{k|k-1} - P_{k|k-1} C_k^T \left(C_k P_{k|k-1} C_k^T + R_k\right)^{-1} C_k P_{k|k-1}$$

Here, for brevity, we introduced the shorthand $P_{i|j} \triangleq P(i|j)$

*The information form of the KF, aka, information filter

- define

$$\mathcal{I}\left(i|j\right) = P\left(i|j\right)^{-1}$$

  we have

$$\begin{aligned}
\mathcal{I}\left(k|k\right) &= \mathcal{I}\left(k|k-1\right) + C\left(k\right)^{T} R\left(k\right)^{-1} C\left(k\right) \\
\mathcal{I}\left(k+1|k\right) &= P\left(k+1|k\right)^{-1} = \left(A_k P_{k|k} A_k^T + Q_k\right)^{-1} \\
&= Q_k^{-1} - Q_k^{-1} A_k \left(A_k^T Q_k^{-1} A_k + \mathcal{I}_{k|k}^{-1}\right)^{-1} A_k^T Q_k^{-1}
\end{aligned}$$

*The information form of the KF, aka, information filter

- and the KF gain is

$$
\begin{aligned}
K(k) &= P_{k|k-1} C_k^T \left[C_k P_{k|k-1} C_k^T + R_k \right]^{-1} \\
&= \mathcal{I}_{k|k-1}^{-1} C_k^T \left[R_k^{-1} - R_k^{-1} C_k \underbrace{\left(C_k^T R_k^{-1} C_k + \underbrace{P_{k|k-1}^{-1}}_{\mathcal{I}_{k|k-1}} \right)^{-1}}_{\mathcal{I}_{k|k}^{-1}} C_k^T R_k^{-1} \right] \\
&= \left[\mathcal{I}_{k|k-1}^{-1} - \mathcal{I}_{k|k-1}^{-1} C_k^T R_k^{-1} C_k \mathcal{I}_{k|k}^{-1} \right] C_k^T R_k^{-1} \\
&= \left[\mathcal{I}_{k|k-1}^{-1} - \mathcal{I}_{k|k-1}^{-1} \left(\mathcal{I}_{k|k} - \mathcal{I}_{k|k-1} \right) \mathcal{I}_{k|k}^{-1} \right] C_k^T R_k^{-1} \\
&= \mathcal{I}_{k|k}^{-1} C_k^T R_k^{-1}
\end{aligned}
$$

*The full set of the information filter

$$
\begin{aligned}
\hat{x}_{k|k-1} &= \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1} \\
\mathcal{I}_{k|k-1} &= Q_{k-1}^{-1} - Q_{k-1}^{-1} A_{k-1} \left(A_{k-1}^T Q_{k-1}^{-1} A_{k-1} + \mathcal{I}_{k-1|k-1}^{-1} \right)^{-1} A_{k-1}^T Q_{k-1}^{-1} \\
\hline
\mathcal{I}_{k|k} &= \mathcal{I}_{k|k-1} + C_k^T R_k^{-1} C_k \\
K_k &= \mathcal{I}_{k|k}^{-1} C_k^T R_k^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k \left[y_k - C_k \hat{x}_{k|k-1} \right]
\end{aligned}
$$

with initial conditions $\hat{x}\,(0|-1) = x_o$ and $\mathcal{I}\,(0|0) = X_0^{-1}$.