7. Kalman Filter and H-inf Control

Karl M.H. Lai

Dept. of EEE, HKU

Kalman Filter

- Fundamental tools for analyzing and solving estimation problems.
- Given a linear discrete-time model:

$$
\begin{aligned}
x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k
\end{aligned}
$$

where

$$
\begin{aligned}
w_k &\sim \mathrm{N}(0, Q_k) \\
v_k &\sim \mathrm{N}(0, R_k) \\
E[w_k w_j^T] &= Q_k \delta_{k-j} \\
E[v_k v_j^T] &= R_k \delta_{k-j} \\
E[v_k w_j^T] &= 0
\end{aligned}
$$

$$
\boxed{\delta_{k-j} = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}}
$$

Goal: Estimate x_k with **noisy** observation $\{y_k\}$

$$
\begin{aligned}
\hat{x}_k^+ &= E[x_k | y_1, y_2, \cdots, y_k] = \textit{a posteriori} \text{ estimate} \\
\hat{x}_k^- &= E[x_k | y_1, y_2, \cdots, y_{k-1}] = \textit{a priori} \text{ estimate}
\end{aligned}
$$

Kalman Filter

- Estimate based on conditioning all past record:

$$\hat{x}_{k|k+N} = E[x_k | y_1, y_2, \cdots, y_k, \cdots, y_{k+N}] = \text{ smoothed estimate}$$

$$\hat{x}_{k|k-M} = E[x_k | y_1, y_2, \cdots, y_{k-M}] = \text{ predicted estimate}$$

- Initial estimate:

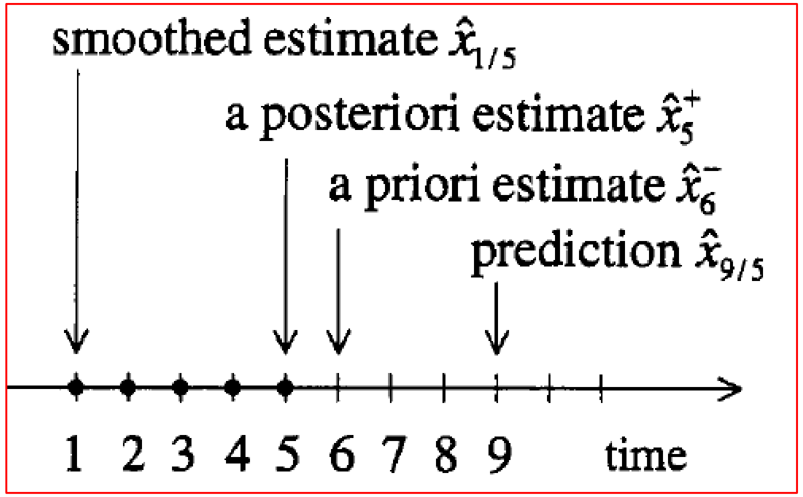
$$\hat{x}_0^+ = E(x_0)$$

- Covariance of Estimation Error:

$$\begin{aligned} P_k^- &= E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T] \\ P_k^+ &= E[(x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^T] \end{aligned}$$

- Mean Propagation:

$$\bar{x}_k = F_{k-1}\bar{x}_{k-1} + G_{k-1}u_{k-1}$$



Kalman Filter

$$\begin{aligned} x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\ y_k &= H_k x_k + v_k \end{aligned}$$

- How to obtain $\hat{x}_1^-$ with $\hat{x}_0^+$? Goal: $\hat{x}_1^- = E(x_1)$ $\bar{x}_k = F_{k-1}\bar{x}_{k-1} + G_{k-1}u_{k-1}$.

$$\hat{x}_1^- = F_0\hat{x}_0^+ + G_0u_0$$

- Time Update Equation: $\hat{x}_k^- = F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1}$
- Uncertainty of Initial Estimation (Covariance):

$$\begin{aligned} P_0^+ &= E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] \\ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \end{aligned}$$

- Covariance Propagation: $P_k = F_{k-1}P_{k-1}F_{k-1}^T + Q_{k-1}$

$$P_1^- = F_0P_0^+F_0^T + Q_0$$

$$P_k^- = F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1}$$

Kalman Filter

$$\begin{aligned} x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\ y_k &= H_k x_k + v_k \end{aligned}$$

- Recall Recursive Least Square Development:

* $\hat{x}_{k-1}$ = estimate before y_k and
$\hat{x}_k$ = estimate after y_k

$$\begin{aligned} K_k &= P_{k-1}H_k^T(H_kP_{k-1}H_k^T + R_k)^{-1} \\ &= P_kH_k^TR_k^{-1} \\ \hat{x}_k &= \hat{x}_{k-1} + K_k(y_k - H_k\hat{x}_{k-1}) \\ P_k &= (I - K_kH_k)P_{k-1}(I - K_kH_k)^T + K_kR_kK_k^T \\ &= (P_{k-1}^{-1} + H_k^TR_k^{-1}H_k)^{-1} \\ &= (I - K_kH_k)P_{k-1} \end{aligned}$$

Least squares estimation		Kalman filtering
$\hat{x}_{k-1}$ = estimate before y_k is processed	$\Longrightarrow$	$\hat{x}_k^-$ = *a priori* estimate
P_{k-1} = covariance before y_k is processed	$\Longrightarrow$	P_k^- = *a priori* covariance
$\hat{x}_k$ = estimate after y_k is processed	$\Longrightarrow$	$\hat{x}_k^+$ = *a posteriori* estimate
P_k = covariance after y_k is processed	$\Longrightarrow$	P_k^+ = *a posteriori* covariance

Recall: Least Square Estimation

$$\frac{\partial J}{\partial \hat{x}} = -y^T H - y^T H + 2\hat{x}^T H^T H$$
$$= 0$$

- Given: $y = Hx + v$
- Cost Function:

$$\begin{aligned} J &= (y - H\hat{x})^T (y - H\hat{x}) \\ &= y^T y - \hat{x}^T H^T y - y^T H\hat{x} + \hat{x}^T H^T H\hat{x} \end{aligned}$$

$$\begin{aligned} H^T y &= H^T H\hat{x} \\ \hat{x} &= (H^T H)^{-1} H^T y \end{aligned}$$ (Estimate based on observations y)

- Weighted Estimation: $E(v_i^2) = \sigma_i^2$

$$\begin{aligned} J &= \epsilon_y^T R^{-1} \epsilon_y \\ &= (y - H\hat{x})^T R^{-1} (y - H\hat{x}) \\ &= y^T R^{-1} y - \hat{x}^T H^T R^{-1} y - y^T R^{-1} H\hat{x} + \hat{x}^T H^T R^{-1} H\hat{x} \end{aligned}$$

$$\begin{aligned} \frac{\partial J}{\partial \hat{x}} &= -y^T R^{-1} H + \hat{x}^T H^T R^{-1} H \\ &= 0 \\ H^T R^{-1} y &= H^T R^{-1} H\hat{x} \end{aligned}$$

$$\boxed{\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} y}$$

- Linear Recursive Estimation:

$$\begin{aligned} y_k &= H_k x + v_k \\ \hat{x}_k &= \hat{x}_{k-1} + K_k (y_k - H_k \hat{x}_{k-1}) \end{aligned}$$

correction term

Recall: Least Square Estimation

$$\boxed{\begin{aligned} y_k &= H_k x + v_k \\ \hat{x}_k &= \hat{x}_{k-1} + K_k(y_k - H_k\hat{x}_{k-1}) \end{aligned}}$$

- Linear Recursive Estimation: (cont')

Estimate Error Mean:

$$\begin{aligned} E(\epsilon_{x,k}) &= E(x - \hat{x}_k) \\ &= E[x - \hat{x}_{k-1} - K_k(y_k - H_k\hat{x}_{k-1})] \\ &= E[\epsilon_{x,k-1} - K_k(H_k x + v_k - H_k\hat{x}_{k-1})] \\ &= E[\epsilon_{x,k-1} - K_k H_k(x - \hat{x}_{k-1}) - K_k v_k] \\ &= (I - K_k H_k)E(\epsilon_{x,k-1}) - K_k E(v_k) \end{aligned}$$

Minimize sum of Var(estimation error):

$$\begin{aligned} J_k &= E[(x_1 - \hat{x}_1)^2)] + \cdots + E[(x_n - \hat{x}_n)^2)] \\ &= E\left(\epsilon_{x1,k}^2 + \cdots + \epsilon_{xn,k}^2\right) \\ &= E\left(\epsilon_{x,k}^T \epsilon_{x,k}\right) \\ &= E\left[\mathrm{Tr}(\epsilon_{x,k}\epsilon_{x,k}^T)\right] \\ &= \mathrm{Tr} P_k \end{aligned}$$

P_k = estimation error covariance

$$\begin{aligned} P_k &= E(\epsilon_{x,k}\epsilon_{x,k}^T) \\ &= E\left\{[(I - K_k H_k)\epsilon_{x,k-1} - K_k v_k][\cdots]^T\right\} \\ &= (I - K_k H_k)E(\epsilon_{x,k-1}\epsilon_{x,k-1}^T)(I - K_k H_k)^T - \\ &\quad K_k E(v_k \epsilon_{x,k-1}^T)(I - K_k H_k)^T - (I - K_k H_k)E(\epsilon_{x,k-1} v_k^T)K_k^T + \\ &\quad K_k E(v_k v_k^T)K_k^T \end{aligned}$$

$$\begin{aligned} E(v_k \epsilon_{x,k-1}^T) &= E(v_k)E(\epsilon_{x,k-1}) \\ &= 0 \end{aligned}$$

* Positive Definite

$$\boxed{P_k = (I - K_k H_k)P_{k-1}(I - K_k H_k)^T + K_k R_k K_k^T}$$

$$\boxed{K_k = P_{k-1}H_k^T(H_k P_{k-1} H_k^T + R_k)^{-1}}$$

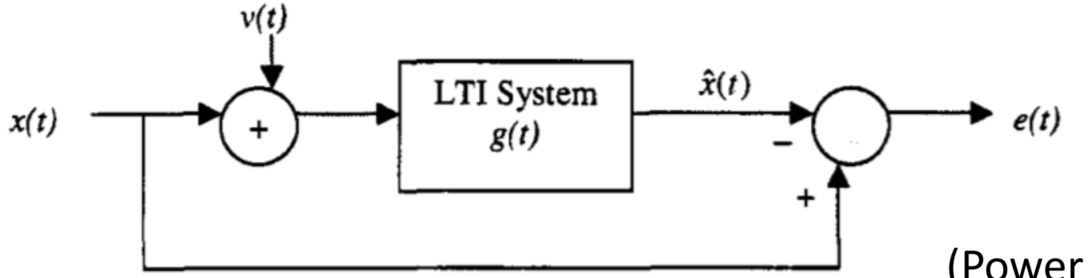
Recall: Least Square Estimation

- General Recursive Least Square Estimation:

$$\begin{aligned} y_k &= H_k x + v_k & \hat{x}_0 &= E(x) \\ x &= \text{constant} & P_0 &= E[(x-\hat{x}_0)(x-\hat{x}_0)^T] \\ E(v_k) &= 0 \\ E(v_k v_i^T) &= R_k \delta_{k-i} \end{aligned}$$

$$\begin{aligned} K_k &= P_{k-1}H_k^T(H_kP_{k-1}H_k^T + R_k)^{-1} \\ &= P_kH_k^TR_k^{-1} \\ \hat{x}_k &= \hat{x}_{k-1} + K_k(y_k - H_k\hat{x}_{k-1}) \\ P_k &= (I-K_kH_k)P_{k-1}(I-K_kH_k)^T + K_kR_kK_k^T \\ &= (P_{k-1}^{-1} + H_k^TR_k^{-1}H_k)^{-1} \\ &= (I-K_kH_k)P_{k-1} \end{aligned}$$

- Wiener Filter:



$$\begin{aligned} \hat{x}(t) &= g(t) * [x(t) + v(t)] \\ \hat{X}(\omega) &= G(\omega)[X(\omega) + V(\omega)] \\ E(\omega) &= X(\omega) - \hat{X}(\omega) \\ &= X(\omega) - G(\omega)[X(\omega) + V(\omega)] \\ &= [1 - G(\omega)]X(\omega) - G(\omega)V(\omega) \end{aligned}$$

(Power Signal)

$$S_e(\omega) = [1 - G(\omega)][1 - G(-\omega)]S_x(\omega) - G(\omega)G(-\omega)S_v(\omega)$$

- Noncasual Filter:

$$\begin{aligned} R_x(\tau) &= \int g(u)[R_x(u-\tau) + R_v(u-\tau)]\,du \\ &= g(\tau) * [R_x(\tau) + R_v(\tau)] \\ S_x(\omega) &= G(\omega)[S_x(\omega) + S_v(\omega)] \end{aligned} \qquad G(\omega) = \frac{S_x(\omega)}{S_x(\omega) + S_v(\omega)}$$

Kalman Filter

- Measurement Update Equations:

$$
\begin{aligned}
K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \\
&= P_k^+ H_k^T R_k^{-1} \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k (y_k - H_k \hat{x}_k^-) \\
P_k^+ &= (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k R_k K_k^T \\
&= \left[(P_k^-)^{-1} + H_k^T R_k^{-1} H_k \right]^{-1} \\
&= (I - K_k H_k) P_k^-
\end{aligned}
$$

- One-Step Kalman Filter with Discrete Riccati:

$$\hat{x}_{k+1}^- = F_k \hat{x}_k^+ + G_k u_k$$ (priori)

$$P_{k+1}^- = F_k \boxed{P_k^+} F_k^T + Q_k$$ Substitute P_k^+

$$
\begin{aligned}
\hat{x}_{k+1}^- &= F_k \left[\hat{x}_k^- + K_k (y_k - H_k \hat{x}_k^-) \right] + G_k u_k \\
&= F_k (I - K_k H_k) \hat{x}_k^- + F_k K_k y_k + G_k u_k
\end{aligned}
$$

(posteriori)

Kalman Filter (Summary 1)

- Discrete KF summary:

1. The dynamic system is given by the following equations:

$$\begin{aligned} x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\ y_k &= H_k x_k + v_k \\ E(w_k w_j^T) &= Q_k \delta_{k-j} \\ E(v_k v_j^T) &= R_k \delta_{k-j} \\ E(w_k v_j^T) &= 0 \end{aligned} \tag{5.17}$$

2. The Kalman filter is initialized as follows:

$$\begin{aligned} \hat{x}_0^+ &= E(x_0) \\ P_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \end{aligned} \tag{5.18}$$

3. The Kalman filter is given by the following equations, which are computed for each time step $k = 1, 2, \cdots$:

$$\begin{aligned} P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1} \\ K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} \\ &= P_k^+ H_k^T R_k^{-1} \\ \hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} = \textit{a priori} \text{ state estimate} \\ \hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H_k\hat{x}_k^-) = \textit{a posteriori} \text{ state estimate} \\ P_k^+ &= (I - K_kH_k)P_k^-(I - K_kH_k)^T + K_kR_kK_k^T \\ &= \left[(P_k^-)^{-1} + H_k^T R_k^{-1} H_k\right]^{-1} \\ &= (I - K_kH_k)P_k^- \end{aligned} \tag{5.19}$$

Kalman Filter:

- Riccati Equation:

$$\begin{aligned} P_{k+1}^- &= F_k(P_k^- - K_k H_k P_k^-)F_k^T + Q_k \\ &= F_k P_k^- F_k^T - F_k K_k H_k P_k^- F_k^T + Q_k \\ &= F_k P_k^- F_k^T - F_k P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- F_k^T + Q_k \end{aligned}$$

(priori based)

$$\begin{aligned} \hat{x}_k^+ &= (I - K_k H_k)(F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1}) + K_k y_k \\ P_k^+ &= (I - K_k H_k)(F_{k-1}P_{k-1}^+ F_{k-1}^T + Q_{k-1}) \end{aligned}$$

(posteriori based)

- Multiple-State System and Steady State KF:

$$\begin{aligned} P_k^- &= F_{k-1}P_{k-1}^+ F_{k-1}^T + Q_{k-1} \\ P_k^+ &= P_k^- - P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{-1} H_k P_k^- \end{aligned}$$

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} (F_k + Q_k F_k^{-T} H_k^T R_k^{-1} H_k) & Q_k F_k^{-T} \\ F_k^{-T} H_k^T R_k^{-1} H_k & F_k^{-T} \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix}$$

$\boxed{\textit{Factorize } P_k^- = A_k B_k^{-1}}$

$A_{k+1}B_{k+1}^{-1} = P_{k+1}^-$ (State w/o proof)

Kalman Filter:

- Multiple State System and Steady State KF:

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} (F + QF^{-T}H^TR^{-1}H) & QF^{-T} \\ F^{-T}H^TR^{-1}H & F^{-T} \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix}$$

$$= \Psi \begin{bmatrix} A_k \\ B_k \end{bmatrix}$$

$$\begin{bmatrix} A_k \\ B_k \end{bmatrix} = \Psi^{k-1} \begin{bmatrix} P_1^- \\ I \end{bmatrix}$$

$$\boxed{\begin{bmatrix} A_\infty \\ B_\infty \end{bmatrix} \approx \Psi^{2^p} \begin{bmatrix} P_1^- \\ I \end{bmatrix} \quad \text{for large } p}$$

- For Scalar System:

$$\begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} = \begin{bmatrix} F + \frac{H^2Q}{FR} & \frac{Q}{F} \\ \frac{H^2}{FR} & \frac{1}{F} \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix}$$

$$= \Psi \begin{bmatrix} A_k \\ B_k \end{bmatrix}$$

$$\begin{bmatrix} A_k \\ B_k \end{bmatrix} = \Psi^{k-1} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix}$$

$$= M \begin{bmatrix} \lambda_1^{k-1} & 0 \\ 0 & \lambda_2^{k-1} \end{bmatrix} M^{-1} \begin{bmatrix} P_1^- \\ 1 \end{bmatrix}$$

(eigendecomposition)

Kalman Filter:

- Scalar System (cont')

$$
\begin{aligned}
\begin{bmatrix} A_k \\ B_k \end{bmatrix} &= \Psi^{k-1} \begin{bmatrix} A_1 \\ B_1 \end{bmatrix} \\
&= M \begin{bmatrix} \lambda_1^{k-1} & 0 \\ 0 & \lambda_2^{k-1} \end{bmatrix} M^{-1} \begin{bmatrix} P_1^- \\ 1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
P_k^- &= \frac{\tau_1 \mu_1^{k-1}(2RH^2P_1^- - \tau_2) - \tau_2 \mu_2^{k-1}(2H^2P_1^- - \tau_1)}{2H^2\mu_1^{k-1}(2RH^2P_1^- - \tau_2) - 2H^2\mu_2^{k-1}(2H^2P_1^- - \tau_1)} \\
\lambda_1 &= \frac{H^2Q + R(F^2+1) + \sigma}{2FR} \\
\lambda_2 &= \frac{H^2Q + R(F^2+1) - \sigma}{2FR} \\
\sigma &= \sqrt{H^2Q + R(F+1)^2}\sqrt{H^2Q + R(F-1)^2} \\
\tau_1 &= H^2Q + R(F^2-1) + \sigma \\
\tau_2 &= H^2Q + R(F^2-1) - \sigma \\
\mu_1 &= H^2Q + R(F^2+1) + \sigma \\
\mu_2 &= H^2Q + R(F^2+1) - \sigma
\end{aligned}
$$

** Time varying KF for scalar invariance system*

$$
\begin{aligned}
M &= \begin{bmatrix} \frac{\tau_1}{2H^2} & \frac{\tau_2}{2H^2} \\ 1 & 1 \end{bmatrix} \\
M^{-1} &= \frac{1}{\tau_1(R-1) + 2\sigma} \begin{bmatrix} 2RH^2 & -\tau_1 \\ -2RH^2 & R\tau_1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\lim_{k\to\infty} P_k^- &= \lim_{k\to\infty} \frac{\tau_1 \mu_1^{k-1}(2RH^2P_1^- - \tau_2) - \tau_2 \mu_2^{k-1}(2H^2P_1^- - \tau_1)}{2H^2\mu_1^{k-1}(2RH^2P_1^- - \tau_2) - 2H^2\mu_2^{k-1}(2H^2P_1^- - \tau_1)} \\
&= \lim_{k\to\infty} \frac{\tau_1 \mu_1^{k-1}(2RH^2P_1^- - \tau_2)}{2H^2\mu_1^{k-1}(2RH^2P_1^- - \tau_2)} \\
&= \boxed{\frac{\tau_1}{2H^2}}
\end{aligned}
$$

Kalman Filter:

- Recall the assumption of KF:
  1. F, Q, H, and R matrices are exactly known
  2. $\{w_k\}$ and $\{v_k\}$ are pure white, zero-mean, and uncorrelated.

$\downarrow$

divergent due to mismodelling

- Strategies to improve KF performance:
  1. Increase arithmetic precision
  2. Use some form of square root filtering
  3. Symmetrize P at each time step: $P = (P + P^T)/2$
  4. Initialize P appropriately to avoid large changes in P
  5. Use a fading-memory filter
  6. Use fictitious process noise (especially for estimating "constants")

Sequential Kalman Filter

$$
\begin{aligned}
y_k &= H_k x_k + v_k \\
K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + R_k)^{\boxed{-1}} \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k (y_k - H_k \hat{x}_k^-) \\
P_k^+ &= (I - K_k H_k) P_k^-
\end{aligned}
$$

Denote $y_k(i) = y_{ik}$

- To implement KF w/o matrix inversion:

1. The system and measurement equations are given as

$$
\begin{aligned}
x_k &= F_{k-1} x_{k-1} + G_{k-1} u_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k \\
w_k &\sim (0, Q_k) \\
v_k &\sim (0, R_k)
\end{aligned} \tag{6.6}
$$

where w_k and v_k are uncorrelated white noise sequences. The measurement covariance R_k is a diagonal matrix given as

$$R_k = \mathrm{diag}(R_{1k}, \cdots, R_{rk}) \tag{6.7}$$

2. The filter is initialized as

$$
\begin{aligned}
\hat{x}_0^+ &= E(x_0) \\
P_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T]
\end{aligned} \tag{6.8}
$$

Sequential Kalman Filter

3. At each time step k, the time-update equations are given as

$$\begin{aligned} P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1} \\ \hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} \end{aligned} \tag{6.9}$$

   This is the same as the standard Kalman filter.

4. At each time step k, the measurement-update equations are given as follows.

   (a) Initialize the *a posteriori* estimate and covariance as

$$\begin{aligned} \hat{x}_{0k}^+ &= \hat{x}_k^- \\ P_{0k}^+ &= P_k^- \end{aligned} \tag{6.10}$$

   These are the *a posteriori* estimate and covariance at time k after zero measurements have been processed; that is, they are equal to the *a priori* estimate and covariance.

Sequential Kalman Filter

4 (b) For $i = 1, \cdots, r$ (where r is the number of measurements), perform the following:

* SKF only makes sense when either R_k is diagonals or R is constant

$$
\begin{aligned}
K_{ik} &= \frac{P^+_{i-1,k} H^T_{ik}}{H_{ik} P^+_{i-1,k} H^T_{ik} + R_{ik}} \\
&= \frac{P^+_{ik} H^T_{ik}}{R_{ik}} \\
\hat{x}^+_{ik} &= \hat{x}^+_{i-1,k} + K_{ik}(y_{ik} - H_{ik}\hat{x}^+_{i-1,k}) \\
P^+_{ik} &= (I - K_{ik}H_{ik}) P^+_{i-1,k} (I - K_{ik}H_{ik})^T + K_{ik} R_{ik} K^T_{ik} \\
&= \left[(P^+_{i-1,k})^{-1} + H^T_{ik} H_{ik} / R_{ik} \right]^{-1} \\
&= (I - K_{ik}H_{ik}) P^+_{i-1,k}
\end{aligned}
\tag{6.11}
$$

(c) Assign the *a posteriori* estimate and covariance as

$$
\begin{aligned}
\hat{x}^+_k &= \hat{x}^+_{rk} \\
P^+_k &= P^+_{rk}
\end{aligned}
\tag{6.12}
$$

Information Filter

$$\boxed{(A + BD^{-1}C)^{-1} = A^{-1} - A^{-1}B(D + CA^{-1}B)^{-1}CA^{-1}}$$

- Propagate P^{-1} instead of P. $P = E[(x - \hat{x})(x - \hat{x})^T]$

P = 0: Perfect information
P = inf: no information

- Define information matrix: $\mathcal{I} = P^{-1}$
- Proved before:

$$(P_k^+)^{-1} = (P_k^-)^{-1} + H_k^T R_k^{-1} H_k$$
$$P_k^- = F_{k-1} P_{k-1}^+ F_{k-1}^T + Q_{k-1}$$

$\rightarrow$

$$\mathcal{I}_k^+ = \mathcal{I}_k^- + H_k^T R_k^{-1} H_k$$
$$\mathcal{I}_k^- = [F_{k-1}(\mathcal{I}_{k-1}^+)^{-1} F_{k-1}^T + Q_{k-1}]^{-1}$$

$$\mathcal{I}_k^- = Q_{k-1}^{-1} - Q_{k-1}^{-1} F_{k-1} (\mathcal{I}_{k-1}^+ + F_{k-1}^T Q_{k-1}^{-1} F_{k-1})^{-1} F_{k-1}^T Q_{k-1}^{-1}$$

- Why P^{-1} instead of P?
  P requires r x r inversion (r = # of measurement), P^{-1} requires n x n inversion (n = # of states) → if r >> n, information filter is much better!

Information Filter

1. The dynamic system is given by the following equations:

$$
\begin{aligned}
x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k \\
w_k &\sim (0, Q_k) \\
v_k &\sim (0, R_k) \\
E(w_k w_j^T) &= Q_k \delta_{k-j} \\
E(v_k v_j^T) &= R_k \delta_{k-j} \\
E(w_k v_k^T) &= 0
\end{aligned}
$$

2. The Kalman filter is initialized as follows:

$$
\begin{aligned}
\hat{x}_0^+ &= E(x_0) \\
\mathcal{I}_0^+ &= \left\{ E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \right\}^{-1}
\end{aligned}
\tag{6.32}
$$

3. The information filter is given by the following equations, which are computed for each time step $k = 1, 2, \cdots$:

3. (cont')

$$
\begin{aligned}
\mathcal{I}_k^- &= Q_{k-1}^{-1} - Q_{k-1}^{-1}F_{k-1}(\mathcal{I}_{k-1}^+ + F_{k-1}^T Q_{k-1}^{-1} F_{k-1})^{-1} F_{k-1}^T Q_{k-1}^{-1} \\
\mathcal{I}_k^+ &= \mathcal{I}_k^- + H_k^T R_k^{-1} H_k \\
K_k &= (\mathcal{I}_k^+)^{-1} H_k^T R_k^{-1} \\
\hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H_k \hat{x}_k^-)
\end{aligned}
$$

Square-Root Update Equation

- Riccati Equation Solution P_k is theoretically positive symmetric semi-definite, but numerical implementation leads to either indefinite or asymmetric. Hence, square-root filtering, by Cholesky Factorization $P = S\,S^T$, is needed to improve the accuracy and overcome the two problems.

The Cholesky Matrix Square Root Algorithm {
For $i = 1, \cdots, n$
{
$S_{ii} = \sqrt{P_{ii} - \sum_{j=1}^{i-1} S_{ij}^2}$
For $j = 1, \cdots, n$
{
$S_{ji} = 0 \quad j < i$
$S_{ji} = \frac{1}{S_{ii}}\left(P_{ji} - \sum_{k=1}^{i-1} S_{jk}S_{ik}\right) \quad j > i$
}
}
}

Consider an n-state discrete LTI system:

$$\begin{aligned} x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\ E(w_k w_k^T) &= Q_k \end{aligned}$$

We can find an orthogonal T, such that

$$\begin{aligned} \begin{bmatrix} (S_k^-)^T \\ 0 \end{bmatrix} &= T \begin{bmatrix} (S_{k-1}^+)^T F_{k-1}^T \\ Q_{k-1}^{T/2} \end{bmatrix} \\ &= \begin{bmatrix} T_1 & T_2 \end{bmatrix} \begin{bmatrix} (S_{k-1}^+)^T F_{k-1}^T \\ Q_{k-1}^{T/2} \end{bmatrix} \end{aligned}$$

Square Root Update Equation

$$T^T T = \begin{bmatrix} T_1^T \\ T_2^T \end{bmatrix} \begin{bmatrix} T_1 & T_2 \end{bmatrix} = \begin{bmatrix} T_1^T T_1 & T_1^T T_2 \\ T_2^T T_1 & T_2^T T_2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

- Since T is orthogonal, $T^T T = I$.

$$\begin{bmatrix} S_k^- & 0 \end{bmatrix} \begin{bmatrix} (S_k^-)^T \\ 0 \end{bmatrix} = \left[T_1 (S_{k-1}^+)^T F_{k-1}^T + T_2 Q_{k-1}^{T/2}\right]^T \left[\cdots\right]$$

$$\begin{aligned} S_k^- (S_k^-)^T &= F_{k-1} S_{k-1}^+ T_1^T T_1 (S_{k-1}^+)^T F_{k-1}^T + Q_{k-1}^{1/2} T_2^T T_2 Q_{k-1}^{T/2} \\ &= F_{k-1} S_{k-1}^+ (S_{k-1}^+)^T F_{k-1}^T + Q_{k-1}^{1/2} Q_{k-1}^{T/2} \end{aligned}$$

- If S_{k-1}^+ is the square root of P_{k-1}^+, $\quad P_k^- = F_{k-1} P_{k-1}^+ F_{k-1}^T + Q_{k-1}$ (time update equation of KF)
- As S and T are not unique, Potter's Algorithm is suggested.

$$P_k^+ = (I - K_k H_k) P_k^- \qquad K_{ik} = \frac{P_{i-1,k}^+ H_{ik}^T}{H_{ik} P_{i-1,k}^+ H_{ik}^T + R_{ik}} \qquad P_{i-1,k}^+ = S_{i-1,k}^+ S_{i-1,k}^{+T}$$

$$P_{0k}^+ = P_k^-$$

$$P_{ik}^+ = (I - K_{ik} H_{ik}) P_{i-1,k}^+$$

$$K_{ik} = \frac{S_{i-1,k}^+ S_{i-1,k}^{+T} H_{ik}^T}{H_{ik} S_{i-1,k}^+ S_{i-1,k}^{+T} H_{ik}^T + R_{ik}}$$

Square Root Update Equation

$$
\begin{aligned}
P_{ik}^{+} &= \left(I - \frac{S_{i-1,k}^{+} S_{i-1,k}^{+T} H_{ik}^{T} H_{ik}}{H_{ik} S_{i-1,k}^{+} S_{i-1,k}^{+T} H_{ik}^{T} + R_{ik}}\right) S_{i-1,k}^{+} S_{i-1,k}^{+T} \\
&= S_{i-1,k}^{+}(I - a\phi\phi^{T}) S_{i-1,k}^{+T}
\end{aligned}
$$

$$
\begin{aligned}
\phi &= S_{i-1,k}^{+T} H_{ik}^{T} \\
a &= \frac{1}{\phi^{T}\phi + R_{ik}}
\end{aligned}
$$

$$I - a\phi\phi^{T} = (I - a\gamma\phi\phi^{T})^{2}$$

$$\gamma = \frac{1}{1 \pm \sqrt{aR_{ik}}}$$

$$S_{ik}^{+} = S_{i-1,k}^{+}(I - a\gamma\phi\phi^{T})$$

1. After the *a priori* covariance square root S_k^- and the *a priori* state estimate $\hat{x}_k^-$ have been computed, initialize

$$
\begin{aligned}
\hat{x}_{0k}^{+} &= \hat{x}_k^{-} \\
S_{0k}^{+} &= S_k^{-} \qquad (6.73)
\end{aligned}
$$

2. For $i = 1, \cdots, r$ (where r is the number of measurements), perform the following.
   (a) Define H_{ik} as the ith row of H_k, y_{ik} as the ith element of y_k, and R_{ik} as the variance of the ith measurement (assuming that R_k is diagonal).
   (b) Perform the following to find the square root of the covariance after the ith measurement has been processed:

$$
\begin{aligned}
\phi_i &= S_{i-1,k}^{+T} H_{ik}^{T} \\
a_i &= \frac{1}{\phi_i^{T}\phi_i + R_{ik}} \\
\gamma_i &= \frac{1}{1 \pm \sqrt{a_i R_{ik}}} \\
S_{ik}^{+} &= S_{i-1,k}^{+}(I - a_i\gamma_i\phi_i\phi_i^{T})
\end{aligned}
$$

Square Root Update Equation

(c) Compute the Kalman gain for the ith measurement as

$$K_{ik} = a_i S_{ik}^{+} \phi_i \tag{6.75}$$

(d) Compute the state estimate update due to the ith measurement as

$$\hat{x}_{ik}^{+} = \hat{x}_{i-1,k}^{+} + K_{ik}(y_{ik} - H_{ik}\hat{x}_{i-1,k}^{+}) \tag{6.76}$$

3. Set the *a posteriori* covariance square root and the *a posteriori* state estimate as

$$\begin{aligned} S_k^{+} &= S_{rk}^{+} \\ \hat{x}_k^{+} &= \hat{x}_{rk}^{+} \end{aligned} \tag{6.77}$$

Another algorithm (triangularization) assumes:

$$\begin{bmatrix} (R_k + H_k P_k^{-} H_k^T)^{T/2} & \tilde{K}_k^T \\ 0 & (S_k^{+})^T \end{bmatrix} = \tilde{T} \begin{bmatrix} R_k^{T/2} & 0 \\ (S_k^{-})^T H_k^T & (S_k^{-})^T \end{bmatrix}$$

$$\begin{bmatrix} (R_k + H_k P_k^{-} H_k^T)^{T/2} & \tilde{K}_k^T \\ 0 & (n \times n \text{ matrix}) \end{bmatrix} = \tilde{T} \begin{bmatrix} R_k^{T/2} & 0 \\ (S_k^{-})^T H_k^T & (S_k^{-})^T \end{bmatrix}$$

with normal Kalman Gain K_k:

$$\tilde{K}_k = K_k (R_k + H_k P_k^{-} H_k^T)^{T/2}$$

$$\boxed{S_k^{+}(S_k^{+})^T = P_k^{-} - K_k H_k P_k^{-}}$$

The goal is to find (n x n matrix) = $\sqrt{P_k^{+}}^T$

U-D Filtering and update equations

- U-D Factorization is computationally easier.

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{bmatrix} = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ u_{12} & 1 & 0 \\ u_{13} & u_{23} & 1 \end{bmatrix}$$

- Consider measurement update equation,

$$P^+ = P^- - P^- H^T (H P^- H^T + R)^{-1} H P^-$$

$$P_i = P_{i-1} - P_{i-1} H_i^T (H_i P_{i-1} H_i^T + R_i)^{-1} H_i P_{i-1}$$

$$\begin{aligned} U_i D_i U_i^T &= U_{i-1} D_{i-1} U_{i-1}^T - \frac{1}{\alpha_i} U_{i-1} D_{i-1} U_{i-1}^T H_i^T H_i U_{i-1} D_{i-1} U_{i-1}^T \\ &= U_{i-1} \left[D_{i-1} - \frac{1}{\alpha_i} (D_{i-1} U_{i-1}^T H_i^T)(D_{i-1} U_{i-1}^T H_i^T)^T \right] U_{i-1}^T \end{aligned}$$

$$\boxed{\bar{U} \bar{D} \bar{U}^T = \left[D_{i-1} - \frac{1}{\alpha_i} (D_{i-1} U_{i-1}^T H_i^T)(D_{i-1} U_{i-1}^T H_i^T)^T \right]}$$

Positive symmetric definite

$$\begin{aligned} U_i D_i U_i^T &= U_{i-1} \bar{U} \bar{D} \bar{U}^T U_{i-1}^T \\ &= (U_{i-1} \bar{U}) \bar{D} (U_{i-1} \bar{U})^T \end{aligned}$$

$$\boxed{\begin{aligned} U_i &= U_{i-1} \bar{U} \\ D_i &= \bar{D} \end{aligned}}$$

U-D Filtering and update equations

- Time Update Equations: $P^- = FP^+F^T + Q$

$$P^- = U^- D^- U^{-T} = FP^+F^T + Q$$

$$P^+ = U^+ D^+ U^{+T}$$

* U^{-T} is not inverse of U, but transpose of U

$$\begin{aligned} P^- &= FP^+F^T + Q \\ &= \begin{bmatrix} FU^+ & I \end{bmatrix} \begin{bmatrix} D^+ & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} U^{+T}F^T \\ I \end{bmatrix} \\ &= W\hat{D}W^T \end{aligned}$$

$$\boxed{U^- D^- U^{-T} = W\hat{D}W^T}$$

$$W^T = \begin{bmatrix} w_1^T & \cdots & w_n^T \end{bmatrix}$$

$$v_k \hat{D} v_j^T = 0 \quad k \neq j \quad \longrightarrow \quad \begin{aligned} v_n &= w_n \\ v_k &= w_k - \sum_{j=k+1}^{n} \frac{w_k \hat{D} v_j^T}{v_j \hat{D} v_j^T} v_j \qquad k = n-1, \cdots, 1 \end{aligned}$$

Define $u(k,j) = \dfrac{w_k \hat{D} v_j^T}{v_j \hat{D} v_j^T} \qquad j,k = 1, \cdots, n$

$$w_k^T = v_k^T + \sum_{j=k+1}^{n} u(k,j) v_j^T$$

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} 1 & u(1,2) & \cdots & u(1,n) \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & u(n-1,n) \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\boxed{W = U^- V}$$

$$\begin{aligned} W\hat{D}W^T &= (U^-V)\hat{D}(U^-V)^T \\ &= U^-(V\hat{D}V^T)U^{-T} \\ &= U^- D^- U^{-T} \end{aligned}$$

General Discrete Time Kalman Filter

$$
\begin{aligned}
x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k \\
w_k &\sim (0, Q_k) \\
v_k &\sim (0, R_k) \\
E[w_k w_j^T] &= Q_k \delta_{k-j} \\
E[v_k v_j^T] &= R_k \delta_{k-j} \\
E[w_k v_j^T] &= M_k \delta_{k-j+1}
\end{aligned}
$$

* Kalman Gain is found by $\min P_k^+$

The Kalman filter is initialized as

$$
\begin{aligned}
\hat{x}_0^+ &= E(x_0) \\
P_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T]
\end{aligned}
\tag{7.13}
$$

For each time step $k = 1, 2, \cdots$, the Kalman filter equations are given as

$$
\begin{aligned}
P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + Q_{k-1} \\
K_k &= (P_k^- H_k^T + M_k)(H_k P_k^- H_k^T + H_k M_k + M_k^T H_k^T + R_k)^{-1} \\
&= P_k^+ \left(H_k^T + (P_k^-)^{-1} M_k\right)\left(R_k - M_k^T (P_k^-)^{-1} M_k\right)^{-1} \\
\hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H_k \hat{x}_k^-) \\
P_k^+ &= (I - K_k H_k) P_k^- (I - K_k H_k)^T + \\
& \quad K_k (H_k M_k + M_k^T H_k^T + R_k) K_k^T - M_k K_k^T - K_k M_k^T \\
&= \Big[(P_k^-)^{-1} + \left(H_k^T + (P_k^-)^{-1} M_k\right)\left(R_k - M_k^T (P_k^-)^{-1} M_k\right)^{-1} \times \\
& \quad \left(H_k + M_k^T (P_k^-)^{-1}\right)\Big]^{-1} \\
&= P_k^- - K_k (H_k P_k^- + M_k^T)
\end{aligned}
\tag{7.14}
$$

Colored Measurement Noise

Define auxiliary signal y_k'.

$$y'_{k-1} = y_k - \psi_{k-1} y_{k-1}$$

$$\begin{aligned}
x_k &= F_{k-1}x_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k \\
v_k &= \psi_{k-1} v_{k-1} + \zeta_{k-1} \\
w_k &\sim (0, Q_k) \\
\zeta_k &\sim (0, Q_{\zeta k}) \\
E[w_k w_j^T] &= Q_k \delta_{k-j} \\
E[\zeta_k \zeta_j^T] &= Q_{\zeta k} \delta_{k-j} \\
E[w_k \zeta_j^T] &= 0
\end{aligned}$$

$$\begin{aligned}
y'_{k-1} &= (H_k x_k + v_k) - \psi_{k-1}(H_{k-1}x_{k-1} + v_{k-1}) \\
&= H_k(F_{k-1}x_{k-1} + w_{k-1}) + v_k - \psi_{k-1}(H_{k-1}x_{k-1} + v_{k-1}) \\
&= (H_k F_{k-1} - \psi_{k-1}H_{k-1})x_{k-1} + H_k w_{k-1} + v_k - \psi_{k-1}v_{k-1} \\
&= (H_k F_{k-1} - \psi_{k-1}H_{k-1})x_{k-1} + (H_k w_{k-1} + \zeta_{k-1}) \\
&= H'_{k-1}x_{k-1} + v'_{k-1}
\end{aligned}$$

$$\begin{aligned}
x_k &= F_{k-1}x_{k-1} + w_{k-1} \\
y'_k &= H'_k x_k + v'_k
\end{aligned}$$

$$\begin{aligned}
E[v'_k {v'_k}^T] &= E[(H_{k+1}w_k + \zeta_k)(w_k^T H_{k+1}^T + \zeta_k^T)] \\
&= H_{k+1}Q_k H_{k+1}^T + Q_{\zeta k} \\
E[w_k {v'_k}^T] &= E[w_k(w_k^T H_{k+1}^T + \zeta_k^T)] \\
&= Q_k H_{k+1}^T
\end{aligned}$$

$$\boxed{K_k = \text{argmin Tr } E\left[(x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^T\right]}$$

$$\boxed{\hat{x}_k^- = E[x_k | y_1, \cdots, y_k]}$$

$$\boxed{\begin{aligned}
\hat{x}_k^+ &= E[x_k | y_1, \cdots, y_{k+1}] \\
&= \hat{x}_k^- + K_k(y'_k - H'_k \hat{x}_k^-)
\end{aligned}}$$

Steady-State Filtering and DARE

* Steady State Filtering is NOT optimal in nature.

$$
\begin{aligned}
x_{k+1} &= x_k + w_k \\
y_k &= x_k + v_k \\
w_k &\sim (0,1) \\
v_k &\sim (0,1)
\end{aligned}
$$

$$
\begin{aligned}
\hat{x}_k^- &= F\hat{x}_{k-1}^+ \\
\hat{x}_k^+ &= \hat{x}_k^- + K_\infty(y_k - H\hat{x}_k^-) \\
&= F\hat{x}_{k-1}^+ + K_\infty(y_k - HF\hat{x}_{k-1}^+) \\
&= (I - K_\infty H)F\hat{x}_{k-1}^+ + K_\infty y_k
\end{aligned}
$$

$$P_{k+1}^- = FP_k^+F^T + Q$$

* DARE has semidefinite solution P_∞ if
  (1) (F, H) is detectable
  (2) (F – MR^{-1} H, G) is stabilizable

* SSKF is stable when eig$((I - K_\infty H)F) < 1$

$$P_{k+1}^- = FP_k^-F^T - FK_kHP_k^-F^T - FK_kM^TF^T + Q$$

$$
\begin{aligned}
P_{k+1}^- = \; & FP_k^-F^T - \\
& F(P_k^-H^T + M)(HP_k^-H^T + HM + M^TH^T + R)^{-1}HP_k^-F^T - \\
& F(P_k^-H^T + M)(HP_k^-H^T + HM + M^TH^T + R)^{-1}M^TF^T + Q \\
= \; & FP_k^-F^T - F(P_k^-H^T + M)(HP_k^-H^T + HM + M^TH^T + R)^{-1} \times \\
& (HP_k^- + M^T)F^T + Q
\end{aligned}
$$

$P_k^- = P_{k+1}^-$ (Steady State Assumption)

$$
\begin{aligned}
P_\infty = \; & FP_\infty F^T - \\
& F(P_\infty H^T + M)(HP_\infty H^T + HM + M^TH^T + R)^{-1} \times \\
& (HP_\infty + M^T)F^T + Q
\end{aligned}
$$

(Discrete Algebraic Riccati Equation)

$$K_\infty = (P_\infty H^T + M)(HP_\infty H^T + HM + M^TH^T + R)^{-1}$$

$$\hat{x}_k^+ = (I - K_\infty H)F\hat{x}_{k-1}^+ + K_\infty y_k$$

Hamiltonian Approach to SSKF

1. Form the Hamiltonian matrix

$$\mathcal{H} = \begin{bmatrix} F^{-T} & F^{-T}H^TR^{-1}H \\ QF^{-T} & F + QF^{-T}H^TR^{-1}H \end{bmatrix} \tag{7.99}$$

   For an n-state Kalman filtering problem, the Hamiltonian matrix will be a $2n \times 2n$ matrix.

2. Compute the eigenvalues of $\mathcal{H}$. If any of them are on the unit circle, then we cannot go any further with this procedure; the Riccati equation does not have a steady-state solution.

3. Collect the n eigenvectors that correspond to the n eigenvalues that are outside the unit circle. Put these n eigenvectors in a matrix partitioned as

$$\begin{bmatrix} \Psi_{12} \\ \Psi_{22} \end{bmatrix} \tag{7.100}$$

   The first column of this matrix is the first eigenvector, the second column is the second eigenvector, etc. Ψ_{12} and Ψ_{22} are both $n \times n$ matrices.

4. Compute the steady-state Riccati equation solution as

$$P_\infty^- = \Psi_{22}\Psi_{12}^{-1} \tag{7.101}$$

   Note that Ψ_{12} must be invertible for this method to work.

Fading Memory Filter

$$
\begin{aligned}
x_k &= F_{k-1}x_{k-1} + G_{k-1}u_{k-1} + w_{k-1} \\
y_k &= H_k x_k + v_k \\
E(w_k w_j^T) &= Q_k \delta_{k-j} \\
E(v_k v_j^T) &= R_k \delta_{k-j} \\
E(w_k v_j^T) &= 0
\end{aligned}
$$

The Kalman filter is initialized as follows:

$$
\begin{aligned}
\hat{x}_0^+ &= E(x_0) \\
\tilde{P}_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T]
\end{aligned}
\tag{7.117}
$$

Choose $\alpha \geq 1$ based on how much you want the filter to forget past measurements. If $\alpha = 1$ then the fading-memory filter is equivalent to the standard Kalman filter. In most applications, α is only slightly greater than 1 (for example, $\alpha \approx 1.01$).

The fading-memory filter is given by the following equations, which are computed for each time step $k = 1, 2, \cdots$:

$$
\begin{aligned}
\tilde{P}_k^- &= \alpha^2 F_{k-1}\tilde{P}_{k-1}^+ F_{k-1}^T + Q_{k-1} \\
K_k &= \tilde{P}_k^- H_k^T (H_k \tilde{P}_k^- H_k^T + R_k)^{-1} \\
&= \tilde{P}_k^+ H_k^T R_k^{-1} \\
\hat{x}_k^- &= F_{k-1}\hat{x}_{k-1}^+ + G_{k-1}u_{k-1} \\
\hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H_k \hat{x}_k^-) \\
\tilde{P}_k^+ &= (I - K_k H_k)\tilde{P}_k^-(I - K_k H_k)^T + K_k R_k K_k^T \\
&= \left[(\tilde{P}_k^-)^{-1} + H_k^T R_k^{-1} H_k\right]^{-1} \\
&= \tilde{P}_k^- - K_k H_k \tilde{P}_k^-
\end{aligned}
\tag{7.118}
$$

* a larger α will lead to a more responsive filter to latest measurement.

* To an extreme, Fading Memory Filter (FMF) can Ignore system modelling and estimate the state solely based on measurement.

* SKF: $\lim_{k\to\infty} K_k = 0$ → ignore latest measurement

Projection Approaches to model KF

- The state is constrained by $Dx = d$ or $Dx \leq d$.
- For perfect measurement (i.e. w/o noise), the system can be modelled as

$$\begin{aligned} x_{k+1} &= F_k x_k + w_k \\ \begin{bmatrix} y_k \\ d \end{bmatrix} &= \begin{bmatrix} H_k \\ D \end{bmatrix} x_k + \begin{bmatrix} v_k \\ 0 \end{bmatrix} \end{aligned}$$

- The constraints can be incorporated into max prob derivation, or project KF estimates into a constrained surface.
- Maximum Probability Approaches: Assume x_0, w_k and v_k are Gaussian.

$$\hat{x}_k = \operatorname{argmax}_{x_k} \operatorname{pdf}(x_k|Y_k) \qquad \operatorname{pdf}(x_k|Y_k) = \frac{\exp[-(x_k - \bar{\bar{x}}_k)^T P_k^{-1}(x_k - \bar{\bar{x}}_k)/2]}{(2\pi)^{n/2}|P_k|^{1/2}}$$

$$\bar{\bar{x}}_k = E(x_k|Y_k)$$ (mean)

$$\tilde{x}_k = \operatorname{argmin}_{\tilde{x}_k}(\tilde{x}_k - \bar{\bar{x}}_k)^T P_k^{-1}(\tilde{x}_k - \bar{\bar{x}}_k) \text{ such that } D\tilde{x}_k = d$$ (solution)

$$\begin{aligned} L &= (\tilde{x}_k - \bar{\bar{x}}_k)^T P_k^{-1}(\tilde{x}_k - \bar{\bar{x}}_k) + 2\lambda^T(D\tilde{x}_k - d) \\ \frac{\partial L}{\partial \tilde{x}} &= P_k^{-1}(\tilde{x}_k - \bar{\bar{x}}_k) + D^T\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= D\tilde{x}_k - d = 0 \end{aligned}$$

(Lagrange Multiplier Approach)

$$\begin{aligned} \lambda &= (DP_kD^T)^{-1}(D\bar{\bar{x}}_k - d) \\ &= (DP_kD^T)^{-1}(D\hat{x}_k - d) \\ \tilde{x}_k &= \bar{\bar{x}}_k - P_kD^T\lambda \\ &= \hat{x}_k - P_kD^T(DP_kD^T)^{-1}(D\hat{x}_k - d) \end{aligned}$$

Projection Approaches to model KF

- Least Square Approaches $\tilde{x} = \text{argmin}_{\tilde{x}} E(||x - \tilde{x}||^2|Y)$ such that $D\tilde{x} = d$

$$E(||x - \tilde{x}||^2|Y) = \int (x - \tilde{x})^T (x - \tilde{x})\,\text{pdf}(x|Y)\,dx$$
$$= \int x^T x\,\text{pdf}(x|Y)\,dx - 2\tilde{x}\int x\,\text{pdf}(x|Y)\,dx + \tilde{x}^T\tilde{x}$$
$$L = E(||x - \tilde{x}||^2|Y) + 2\lambda^T(D\tilde{x} - d)$$
$$= \int x^T x\,\text{pdf}(x|Y)\,dx - 2\tilde{x}\int x\,\text{pdf}(x|Y)\,dx + \tilde{x}^T\tilde{x} + 2\lambda^T(D\tilde{x} - d)$$

$$\hat{x} = E(x|Y) = \int x\,\text{pdf}(x|Y)\,dx$$

$$\frac{\partial L}{\partial \tilde{x}} = -2\hat{x} + 2\tilde{x} + 2D^T\lambda = 0$$
$$\frac{\partial L}{\partial \lambda} = D\tilde{x} - d = 0$$

$$\boxed{\begin{aligned}\lambda &= (DD^T)^{-1}(D\hat{x} - d)\\ \tilde{x} &= \hat{x} - D^T(DD^T)^{-1}(D\hat{x} - d)\end{aligned}}$$

- General Projection: $\tilde{x} = \text{argmin}_{\tilde{x}}(\tilde{x} - \hat{x})^T W(\tilde{x} - \hat{x})$ such that $D\tilde{x} = d$

$$\tilde{x} = \hat{x} - W^{-1}D^T(DW^{-1}D^T)^{-1}(D\hat{x} - d)$$

1. Unbiased estimate
2. W = P^{-1} (min variance)
3. W = I (~ unconstrained)

Continuous Kalman Filter

Conversion:

$$
\begin{aligned}
F &= \exp(AT) \\
&\approx (I + AT) \text{ for small } T \\
G &= (\exp(AT) - I)A^{-1}B \\
&\approx BT \text{ for small } T \\
\Lambda &= (\exp(AT) - I)A^{-1} \\
&\approx IT \text{ for small } T \\
H &= C \\
w_k &\sim (0, Q), \quad Q = Q_c T \\
v_k &\sim N(0, R), \quad R = R_c / T
\end{aligned}
$$

- Consider a continuous system:

$$
\begin{aligned}
\dot{x} &= Ax + Bu + w \\
y &= Cx + v \\
w &\sim (0, Q_c) \\
v &\sim (0, R_c)
\end{aligned}
$$

- Discretize the system:

$$
\begin{aligned}
x_k &= F x_{k-1} + G u_{k-1} + \Lambda w_{k-1} \\
y_k &= H x_k + v_k
\end{aligned}
$$

- Discrete-time Kalman Gain: $K_k = P_k^- H^T (H P_k^- H^T + R)^{-1}$

$$
\begin{aligned}
K_k &= P_k^- C^T (C P_k^- C^T + R_c/T)^{-1} \\
\frac{K_k}{T} &= P_k^- C^T (C P_k^- C^T T + R_c)^{-1} \\
\lim_{T \to 0} \frac{K_k}{T} &= P_k^- C^T R_c^{-1}
\end{aligned}
$$

- Error Covariance:

$$
\begin{aligned}
P_k^+ &= (I - K_k H) P_k^- \\
P_{k+1}^- &= F P_k^+ F^T + Q
\end{aligned}
$$

Continuous Kalman Filter

Result:

$$\boxed{\begin{aligned} \dot{\hat{x}} &= A\hat{x} + Bu + K(y - C\hat{x}) \\ K &= PC^T R_c^{-1} \end{aligned}}$$

- For small T and substitute P_k^+

$$\begin{aligned} P_{k+1}^- &= (I + AT)P_k^+(I + AT)^T + Q_c T \\ &= P_k^+ + (AP_k^+ + P_k^+ A^T + Q_c)T + AP_k^+ A^T T^2 \end{aligned}$$

$$\begin{aligned} P_{k+1}^- &= (I - K_k C)P_k^- + AP_k^+ A^T T^2 + \\ &\quad [A(I - K_k C)P_k^- + (I - K_k C)P_k^- A^T + Q_c]T \end{aligned}$$

- Taking the difference,

$$\begin{aligned} \frac{P_{k+1}^- - P_k^-}{T} &= \frac{-K_k C P_k^-}{T} + AP_k^+ A^T T + \\ &\quad (AP_k^- + AK_k C P_k^- + P_k^- A^T - K_k C P_k^- A^T + Q_c) \end{aligned}$$

$$\boxed{\begin{aligned} \dot{P} &= \lim_{T \to 0} \frac{P_{k+1}^- - P_k^-}{T} \\ &= -PC^T R_c^{-1} CP + AP + PA^T + Q_c \end{aligned}}$$

(measurement update)

$$\begin{aligned} \hat{x}_k^- &= F\hat{x}_{k-1}^+ + Gu_{k-1} \\ \hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - H\hat{x}_k^-) \end{aligned}$$

$$\begin{aligned} \hat{x}_k^+ &= F\hat{x}_{k-1}^+ + Gu_{k-1} + K_k(y_k - HF\hat{x}_{k-1}^+ - HGu_{k-1}) \\ &\approx (I + AT)\hat{x}_{k-1}^+ + BTu_{k-1} + \\ &\quad K_k(y_k - C(I + AT)\hat{x}_{k-1}^+ - CBTu_{k-1}) \end{aligned}$$

Continuous Kalman Filter

The continuous-time Kalman filter can be summarized as follows.

1. The continuous-time system dynamics and measurement equations are given as

$$\begin{aligned}
\dot{x} &= Ax + Bu + w \\
y &= Cx + v \\
w &\sim (0, Q_c) \\
v &\sim (0, R_c)
\end{aligned} \tag{8.36}$$

   Note that $w(t)$ and $v(t)$ are continuous-time white noise processes.

2. The continuous-time Kalman filter equations are given as

$$\begin{aligned}
\hat{x}(0) &= E[x(0)] \\
P(0) &= E[(x(0) - \hat{x}(0))(x(0) - \hat{x}(0))^T] \\
K &= PC^T R_c^{-1} \\
\dot{\hat{x}} &= A\hat{x} + Bu + K(y - C\hat{x}) \\
\dot{P} &= -PC^T R_c^{-1} CP + AP + PA^T + Q_c
\end{aligned} \tag{8.37}$$

Solutions to Riccati Equations

- Transition Matrix Approach:

$$P = \Lambda Y^{-1}$$

$$\boxed{\dot{P} = AP + PA^T - PC^TR^{-1}CP + Q}$$

$$\begin{aligned}\dot{P} &= \dot{\Lambda}Y^{-1} + \Lambda\frac{d}{dt}(Y^{-1}) \\ &= \dot{\Lambda}Y^{-1} - \Lambda Y^{-1}\dot{Y}Y^{-1}\end{aligned}$$

$$\dot{P} = A\Lambda Y^{-1} + \Lambda Y^{-1}A^T - \Lambda Y^{-1}C^TR^{-1}C\Lambda Y^{-1} + Q$$

$$\dot{P}Y = \dot{\Lambda} - \Lambda Y^{-1}\dot{Y}$$

$$\begin{aligned}\dot{\Lambda} - \Lambda Y^{-1}\dot{Y} &= A\Lambda + \Lambda Y^{-1}A^TY - \Lambda Y^{-1}C^TR^{-1}C\Lambda + QY \\ \dot{\Lambda} &= A\Lambda + QY + \Lambda Y^{-1}(\dot{Y} + A^TY - C^TR^{-1}C\Lambda)\end{aligned}$$

Condition for valid factorization:

$$\dot{Y} = C^TR^{-1}C\Lambda - A^TY \qquad \dot{\Lambda} = A\Lambda + QY$$

Initial condition:

$$\begin{aligned}\Lambda(0) &= P(0) \\ Y(0) &= I\end{aligned}$$

$$\begin{aligned}\begin{bmatrix}\dot{\Lambda} \\ \dot{Y}\end{bmatrix} &= \begin{bmatrix}A & Q \\ C^TR^{-1}C & -A^T\end{bmatrix}\begin{bmatrix}\Lambda \\ Y\end{bmatrix} \\ &= J\begin{bmatrix}\Lambda \\ Y\end{bmatrix}\end{aligned}$$

$$\begin{bmatrix}\Lambda(t+T) \\ Y(t+T)\end{bmatrix} = \exp(JT)\begin{bmatrix}\Lambda(t) \\ Y(t)\end{bmatrix}$$

$$\begin{bmatrix}\Lambda(t+T) \\ Y(t+T)\end{bmatrix} = \begin{bmatrix}\phi_{11}(T) & \phi_{12}(T) \\ \phi_{21}(T) & \phi_{22}(T)\end{bmatrix}\begin{bmatrix}\Lambda(t) \\ Y(t)\end{bmatrix}$$

Solutions to Riccati Equations

$$P(t+T)Y(t+T) = \phi_{11}(T)P(t)Y(t) + \phi_{12}(T)Y(t)$$

$$\begin{aligned} P(t+T)\left[\phi_{21}(T)P(t)Y(t) + \phi_{22}(T)Y(t)\right] &= \phi_{11}(T)P(t)Y(t) + \phi_{12}(T)Y(t) \\ P(t+T)\left[\phi_{21}(T)P(t) + \phi_{22}(T)\right] &= \phi_{11}(T)P(t) + \phi_{12}(T) \end{aligned}$$

$$P(t+T) = \left[\phi_{11}(T)P(t) + \phi_{12}(T)\right]\left[\phi_{21}(T)P(t) + \phi_{22}(T)\right]^{-1}$$

This can avoid integrating the Riccati Equations to solve for P.

2. The Chandrasekhar algorithm can be summarized as follows.

   1. Compute $\dot{P}(0)$.
   2. Use the method of Section 8.3.2.2 to find M_1 and M_2 matrices that satisfy $\boxed{\dot{P}(0) = M_1 M_1^T - M_2 M_2^T.}$ *Chandrasekhar Factorization*
   3. Initialize $Y_1(0) = M_1$, $Y_2(0) = M_2$, and $K(0) = P(0)C^T R^{-1}$.
   4. Integrate K, Y_1, and Y_2 as follows:

$$\begin{aligned} \dot{K} &= (Y_1 Y_1^T - Y_2 Y_2^T)C^T R^{-1} \\ \dot{Y}_1 &= (A - KC)Y_1 \\ \dot{Y}_2 &= (A - KC)Y_2 \end{aligned} \qquad (8.90)$$

$$\dot{P}(0) = SDS^T$$

$$M_1 = \begin{bmatrix} S_{11} \\ S_{21} \\ S_{31} \end{bmatrix} \sqrt{D_1}$$

$$M_2 = \begin{bmatrix} S_{12} \\ S_{22} \\ S_{32} \end{bmatrix} \sqrt{D_2}$$

Continuous KF with correlated noise

1. The system dynamics and measurement equation are given as

$$\begin{aligned}\dot{x} &= Ax + w \\ w &\sim (0, Q) \\ y &= Cx + v \\ v &\sim (0, R) \\ E[w(t)v^T(\tau)] &= M\delta(t - \tau)\end{aligned}$$

2. The continuous-time Kalman filter is given as

$$\begin{aligned}\dot{P} &= AP + PA^T + Q - KRK^T \\ K &= (PC^T + M)R^{-1} \\ \dot{\hat{x}} &= A\hat{x} + K(y - C\hat{x})\end{aligned}$$

$$\begin{aligned}\dot{x} &= Ax + w + MR^{-1}(y - Cx - v) \\ &= (A - MR^{-1}C)x + MR^{-1}y + (w - MR^{-1}v) \\ &= \tilde{A}x + \tilde{u} + \tilde{w}\end{aligned}$$

$$\begin{aligned}E(\tilde{w}v^T) &= E[(w - MR^{-1}v)v^T] \\ &= E(wv^T) - MR^{-1}E(vv^T) \\ &= M - M \\ &= 0\end{aligned}$$

$$\begin{aligned}\dot{P} &= \tilde{A}P + P\tilde{A}^T - PC^TR^{-1}CP + \tilde{Q} \\ &= (A - MR^{-1}C)P + P(A - MR^{-1}C)^T - PC^TR^{-1}CP + \\ &\quad Q - MR^{-1}M^T\end{aligned}$$

$$\dot{P} = AP + PA^T + Q - \tilde{K}R\tilde{K}^T$$

$$\begin{aligned}\tilde{K} &= K + MR^{-1} \\ &= PC^TR^{-1} + MR^{-1} \\ &= (PC^T + M)R^{-1}\end{aligned}$$

$$\begin{aligned}\dot{\hat{x}} &= \tilde{A}\hat{x} + \tilde{u} + K(y - C\hat{x}) \\ &= (A - MR^{-1}C)\hat{x} + MR^{-1}y + K(y - C\hat{x}) \\ &= A\hat{x} - MR^{-1}C\hat{x} + MR^{-1}y + (\tilde{K} - MR^{-1})(y - C\hat{x}) \\ &= A\hat{x} + \tilde{K}(y - C\hat{x})\end{aligned}$$

Continuous KF with colored measurement noise

$$\begin{aligned} \dot{x} &= Ax + w \\ w &\sim (0, Q) \\ y &= Cx + v \\ \dot{v} &= Nv + \phi \\ \phi &\sim (0, \Phi) \end{aligned}$$

Make the following matrix definitions:

$$\begin{aligned} \tilde{C} &= \dot{C} + CA - NC \\ \tilde{R} &= CQC^T + \Phi \\ M &= QC^T \end{aligned}$$

Initialize the Kalman filter as

$$\begin{aligned} K(0) &= [P(0)C^T + M]\tilde{R}^{-1} \\ z(0) &= \hat{x}(0) - K(0)y(0) \end{aligned}$$

Integrate P, K, and z using the following equations:

$$\begin{aligned} \dot{P} &= AP + PA^T + Q - K\tilde{R}K^T \\ \dot{K} &= \frac{d}{dt}[(PC^T + M)\tilde{R}^{-1}] \\ \dot{z} &= (A - K\tilde{C})\hat{x} - (\dot{K} + KN)y \end{aligned}$$

Note that K equation can be simplified if Q, C, Φ are constant: $\dot{K} = \dot{P}C^T\tilde{R}^{-1}$

Compute the state estimate as: $\hat{x} = z + Ky$

Wiener Filter is a Kalman Filter

Differential Riccati Equations: $\dot{P} = -PC^T R^{-1} CP + AP + PA^T + Q$

Continuous Algebraic Riccati Equation: (CARE)

$$-PC^T R^{-1} CP + AP + PA^T + Q = 0$$

Steady State Kalman Filter: $K = PC^T R^{-1}$

$$\dot{\hat{x}} = (A - KC)\hat{x} + Ky$$

Consider the steady-state continuous-time Kalman filter.

$$\dot{\hat{x}} = Ax + K(y - C\hat{x})$$

Taking the Laplace transform of both sides of this equation gives

$$\begin{aligned} (sI - A + KC)\hat{X}(s) &= KY(s) \\ \hat{X}(s) &= (sI - A + KC)^{-1} KY(s) \end{aligned}$$

Duality between optimal control and estimation

$$\begin{aligned}\dot{x} &= Ax + w\\ w &\sim N(0,Q)\\ y &= Cx + v\\ v &\sim N(0,R)\end{aligned}$$

Cost Function:

$$J_e = \int_0^{t_f} E[(x-\hat{x})^T(x-\hat{x})]\,dt$$

The optimal estimator (the Kalman filter) is given as

$$\begin{aligned}P_e(0) &= E[(x(0)-\hat{x}(0))(x(0)-\hat{x}(0))^T]\\ \dot{P}_e &= AP_e + P_eA^T - P_eC^TR^{-1}CP_e + Q\\ K_e &= P_eC^TR^{-1}\\ \dot{\hat{x}} &= A\hat{x} + K_e(y - C\hat{x})\end{aligned}$$

The optimal control problem begins with the system

$$\dot{x} = Ax + Cu$$

$$\boxed{J_c = x^T\phi x\big|_{t_f} + \int_0^{t_f}(x^TQx + u^TRu)\,dt}$$

Find u for optimal control

Optimal Controller:

$$\begin{aligned}P_c(t_f) &= \phi(t_f)\\ \dot{P}_c &= -A^TP_c - P_cA + P_cCR^{-1}C^TP_c - Q\\ K_c &= R^{-1}C^TP_c\\ u &= -K_cx\end{aligned}$$

Multiple Model Estimation

- System model changes depending on unknown factors. Multiple KF can be used and the state estimates are combined to obtain a refined result.
- For N mutually exclusive event x_1, x_2, …, x_N,

$$\begin{aligned}
\Pr(y) &= \Pr(y|x_1)\Pr(x_1) + \cdots + \Pr(y|x_N)\Pr(x_N) \\
\Pr(x|y) &= \frac{\mathrm{pdf}(y|x)\Pr(x)}{\sum_{i=1}^{N} \mathrm{pdf}(y|x_i)\Pr(x_i)}
\end{aligned}$$

- Assume a time-invariant system:

$$\begin{aligned}
x_k &= F x_{k-1} + G u_{k-1} + w_{k-1} \\
y_k &= H x_k + v_k \\
w_k &\sim N(0, Q) \\
v_k &\sim N(0, R)
\end{aligned}$$

- Given a measurement y_k, what is the probability $p = p_j$?

$$\Pr(p_j|y_k) = \frac{\mathrm{pdf}(y_k|p_j)\Pr(p_j)}{\sum_{i=1}^{N} \mathrm{pdf}(y_k|p_i)\Pr(p_i)}$$

Multiple Model Estimation

- If the estimate is accurate, $\text{pdf}(y_k|p_j) \approx \text{pdf}(y_k|\hat{x}_k^-)$
- Since $y_k \approx H\hat{x}_k^- + v_k$

$$\begin{aligned} \text{pdf}(y_k|p_j) &\approx \text{pdf}(y_k - H_k\hat{x}_k^-) \\ &= \text{pdf}(r_k) \end{aligned}$$

- With v_k, w_k and x_k are Gaussian, r_k is also Gaussian.

$$\text{pdf}(y_k|p_j) \approx \frac{\exp(-r_k^T S_k^{-1} r_k/2)}{(2\pi)^{q/2}|S_k|^{1/2}} \qquad r_k = y_k - H_k\hat{x}_k^-,\ S_k = H_k P_k^- H_k^T + R_k$$

$$\Pr(p_j|y_{k-1}) = \frac{\Pr(y_{k-1}|p_j)\Pr(p_j)}{\Pr(y_{k-1})} \qquad \Pr(y_{k-1}|p_j) = \Pr(y_{k-1}) = 1$$

$$\Pr(p_j|y_{k-1}) = \Pr(p_j)$$

1. For $j = 1, \cdots, N$, initialize the probabilities of each parameter set before any measurements are obtained. These probabilities are denoted as $\Pr(p_j|y_0)$ $(j = 1, \cdots, N)$.

Multiple Model Estimation

2. At each time step k we perform the following steps.

   (a) Run N Kalman filters, one for each parameter set p_j $(j = 1, \cdots, N)$. The *a priori* state estimate and covariance of the jth filter are denoted as $\hat{x}_{kj}^-$ and P_{kj}^-.

   (b) After the measurement at time k is received, for each parameter set approximate the pdf of y_k given p_j as follows:

$$\mathrm{pdf}(y_k|p_j) \approx \frac{\exp(-r_k^T S_k^{-1} r_k/2)}{(2\pi)^{q/2}|S_k|^{1/2}} \tag{10.28}$$

   where $r_k = y_k - H_k \hat{x}_{kj}^-$, $S_k = H P_{kj}^- H^T + R_k$, and q is the number of measurements.

   (c) Estimate the probability that $p = p_j$ as follows.

$$\Pr(p_j|y_k) = \frac{\mathrm{pdf}(y_k|p_j)\Pr(p_j|y_{k-1})}{\sum_{i=1}^N \mathrm{pdf}(y_k|p_i)\Pr(p_i|y_{k-1})} \tag{10.29}$$

   (d) Now that each parameter set p_j has an associated probability, we can weight each $\hat{x}_{kj}^-$ and P_{kj}^- accordingly to obtain

$$\begin{aligned} \hat{x}_k^- &= \sum_{j=1}^N \Pr(p_j|y_k)\hat{x}_{kj}^- \\ P_k^- &= \sum_{j=1}^N \Pr(p_j|y_k)P_{kj}^- \end{aligned} \tag{10.30}$$

   (e) We can estimate the true parameter set in one of several ways, depending on our application. For example, we can use the parameter set with the highest conditional probability as our parameter estimate, or we can estimate the parameter set as a weighted average of the parameter sets:

$$\hat{p} = \begin{cases} \operatorname{argmax}_{p_j} \Pr(p_j|y_k) & \text{max-probability method} \\ \sum_{j=1}^N \Pr(p_j|y_k)p_j & \text{weighted-average method} \end{cases} \tag{10.31}$$

Multiple Model Estimation

- Consider a second-order system:

$$\begin{aligned} \dot{x} &= Ax + Bw_1 \\ &= \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} x + \begin{bmatrix} 0 \\ \omega_n^2 \end{bmatrix} w_1 \\ y_k &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_k + v_k \\ w_1 &\sim N(0, Q_c) \\ v_k &\sim N(0, R) \end{aligned}$$

- The state equation can be written as:

$$\begin{aligned} \dot{x} &= Ax + w \\ w &\sim N(0, BQ_cB^T) \end{aligned}$$

- Discretize the system:

$$\begin{aligned} x_k &= Fx_{k-1} + \Lambda w'_{k-1} \\ y_k &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_k + v_k \\ F &= \exp(AT) \\ \Lambda &= (F - I)F^{-1} \end{aligned}$$

- Recall covariance noise w_k:

$$Q' \approx BQ_cB^TT$$

- The discrete-time process dynamic is:

$$\begin{aligned} x_k &= Fx_{k-1} + w_{k-1} \\ w_k &\sim N(0, Q) \\ Q &= (F - I)F^{-1}(BQB^TT)F^{-T}(F^T - I) \end{aligned}$$

Reduced Order KF: Schmidt-Kalman Filter

- It can be used if the state can be decoupled from each other in dynamic equation.
- Satellite navigation with colored noise was the motivation of this approach.
- Suppose a system:

$$
\begin{aligned}
\begin{bmatrix} \tilde{x}_{k+1} \\ \tilde{\tilde{x}}_{k+1} \end{bmatrix} &= \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix} \begin{bmatrix} \tilde{x}_k \\ \tilde{\tilde{x}}_k \end{bmatrix} + \begin{bmatrix} \tilde{w}_k \\ \tilde{\tilde{w}}_k \end{bmatrix} \\
\tilde{w}_k &\sim (0, Q_1) \\
\tilde{\tilde{w}}_k &\sim (0, Q_2) \\
y_k &= \begin{bmatrix} H_1 & H_2 \end{bmatrix} \begin{bmatrix} \tilde{x}_k \\ \tilde{\tilde{x}}_k \end{bmatrix} + v_k \\
v_k &\sim (0, R)
\end{aligned}
$$

* We don't care $\tilde{\tilde{x}}_k$

- such that the estimation error covariance can be partitioned as $P = \begin{bmatrix} \tilde{P} & \Sigma \\ \Sigma^T & \tilde{\tilde{P}} \end{bmatrix}$
- Kalman Gain: $K = P^- H^T (H P^- H^T + R)^{-1}$

Reduced Order KF: Schmidt-Kalman Filter

- Kalman Gain: $K = \begin{bmatrix} \tilde{K} \\ \tilde{\tilde{K}} \end{bmatrix}$

$$= \begin{bmatrix} \tilde{P}^- & \Sigma^- \\ (\Sigma^-)^T & \tilde{\tilde{P}}^- \end{bmatrix} \times$$

$$\begin{bmatrix} H_1^T \\ H_2^T \end{bmatrix} \left[\begin{pmatrix} H_1 & H_2 \end{pmatrix} \begin{pmatrix} \tilde{P}^- & \Sigma^- \\ (\Sigma^-)^T & \tilde{\tilde{P}}^- \end{pmatrix} \begin{pmatrix} H_1^T \\ H_2^T \end{pmatrix} + R \right]^{-1}$$

$$\boxed{\tilde{K} = (\tilde{P}^- H_1^T + \Sigma^- H_2^T)\alpha^{-1}}$$

$$\alpha = H_1 \tilde{P}^- H_1^T + H_1 \Sigma^- H_2^T + H_2 (\Sigma^-)^T H_1^T + H_2 \tilde{\tilde{P}}^- H_2^T + R$$

- Measurement Update Equation:

$$\begin{bmatrix} \hat{\tilde{x}}_k^+ \\ \hat{\tilde{\tilde{x}}}_k^+ \end{bmatrix} = \begin{bmatrix} \tilde{K} \\ \tilde{\tilde{K}} \end{bmatrix} \left(y_k - H_1 \hat{\tilde{x}}_k^- - H_2 \hat{\tilde{\tilde{x}}}_k^- \right)$$

$$\hat{x}_k^+ = \hat{x}_k^- + K(y_k - H\hat{x}_k^-)$$

- Set $\hat{\tilde{\tilde{x}}}_k^- = 0$

$$\boxed{\hat{\tilde{x}}_k^+ = \hat{\tilde{x}}_k^- + \tilde{K}\left(y_k - H_1 \hat{\tilde{x}}_k^-\right)}$$

Reduced Order KF: Schmidt-Kalman Filter

- Measurement Update Equation for P: $P^+ = (I-KH)P^-(I- KH)^T + KRK^T$

$$\begin{bmatrix} \tilde{P}^+ & \Sigma^+ \\ (\Sigma^+)^T & \tilde{\tilde{P}}^+ \end{bmatrix} = \left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} \tilde{K} \\ \tilde{\tilde{K}} \end{pmatrix}\begin{pmatrix} H_1 & H_2 \end{pmatrix}\right]\begin{bmatrix} \tilde{P}^- & \Sigma^- \\ (\Sigma^-)^T & \tilde{\tilde{P}}^- \end{bmatrix} \times$$

$$\left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} \tilde{K} \\ \tilde{\tilde{K}} \end{pmatrix}\begin{pmatrix} H_1 & H_2 \end{pmatrix}\right]^T + \begin{pmatrix} \tilde{K} \\ \tilde{\tilde{K}} \end{pmatrix} R \begin{pmatrix} \tilde{K}^T & \tilde{\tilde{K}}^T \end{pmatrix}$$

- Assume $\tilde{\tilde{K}} = 0$ with large $\tilde{\tilde{x}}$ measurement noise or small H_2

$$\begin{aligned} \tilde{P}^+ &= (I - \tilde{K}H_1)\tilde{P}^-(I - \tilde{K}H_1)^T - \tilde{K}H_2(\Sigma^-)^T(I - \tilde{K}H_1)^T - \\ &\quad (I - \tilde{K}H_1)\Sigma^- H_2^T\tilde{K}^T + \tilde{K}H_2\tilde{\tilde{P}}^- H_2^T\tilde{K}^T + \tilde{K}R\tilde{K}^T \end{aligned}$$

$$\begin{aligned} \tilde{P}^+ &= \tilde{P}^- - \tilde{K}H_1\tilde{P}^- - \tilde{P}^- H_1^T\tilde{K}^T + \tilde{K}\alpha\tilde{K}^T - \tilde{K}H_2(\Sigma^-)^T - \Sigma^- H_2^T\tilde{K}^T \\ &= \tilde{P}^- - \tilde{K}H_1\tilde{P}^- - \tilde{P}^- H_1^T\tilde{K}^T + (\tilde{P}^- H_1^T + \Sigma^- H_2^T)\tilde{K}^T - \tilde{K}H_2(\Sigma^-)^T - \\ &\quad \Sigma^- H_2^T\tilde{K}^T \\ &= (I - \tilde{K}H_1)\tilde{P}^- - \tilde{K}H_2(\Sigma^-)^T \end{aligned}$$

$$\begin{aligned} \Sigma^+ &= (I - \tilde{K}H_1)\Sigma^- - \tilde{K}H_2\tilde{\tilde{P}}^- \\ \tilde{\tilde{P}}^+ &= \tilde{\tilde{P}}^- \end{aligned}$$

Robust Kalman Filter

- It is assumed that noise w_k and v_k are Gaussian and system matrix F_k, measurement matrix H_k and noise variance matrix Q_k and R_k are known.
- Here, an adaptive, or robust, KF are introduced for uncertainties in Q_k and R_k.
- With an LTI system:

$$\begin{aligned} x_{k+1} &= Fx_k + w_k \\ y_k &= Hx_k + v_k \\ w_k &\sim (0, Q) \\ v_k &\sim (0, R) \end{aligned}$$

- Assume a steady-state gain K (not necessarily Kalman gain) is used in a predictor/ corrector type of state estimator and the state estimate update equations are as follows.

$$\begin{aligned} \hat{x}_{k+1}^- &= F\hat{x}_k^+ \\ \hat{x}_{k+1}^+ &= \hat{x}_{k+1}^- + K(y_{k+1} - H\hat{x}_{k+1}^-) \\ &= F\hat{x}_k^+ + K(Hx_{k+1} + v_{k+1} - HF\hat{x}_k^+) \\ &= KHx_{k+1} + (I - KH)F\hat{x}_k^+ + Kv_{k+1} \\ &= (KHFx_k + KHw_k) + (I - KH)F\hat{x}_k^+ + Kv_{k+1} \end{aligned}$$

Robust Kalman Filter

- The error in posteriori state estimate:

$$\begin{aligned} e_{k+1} &= x_{k+1} - \hat{x}_{k+1}^{+} \\ &= (Fx_k + w_k) - [(KHFx_k + KHw_k) + (I - KH)F\hat{x}_k^{+} + Kv_{k+1}] \\ &= (I - KH)Fx_k + (I - KH)w_k - (I - KH)F\hat{x}_k^{+} - Kv_{k+1} \\ &= (I - KH)Fe_k + (I - KH)w_k - Kv_{k+1} \end{aligned}$$

- Estimation Error Covariance:

$$\begin{aligned} P_{k+1} &= E(e_{k+1}e_{k+1}^T) \\ &= (I - KH)FP_kF^T(I - KH)^T + (I - KH)Q(I - KH)^T + \\ &\quad KRK^T \end{aligned}$$

- The steady state covariance P should satisfy the Riccati Equation:

$$P = (I - KH)FPF^T(I - KH)^T + (I - KH)Q(I - KH)^T + KRK^T$$

- Note that we haven’t made any assumption on the optimality of K.

Robust Kalman Filter

- What if there is no measurement and process noise? → R = 0
  P_1 = estimation error covariance with R = 0:
  P_2 = estimation error covariance with Q = 0:

$$
\begin{aligned}
P_1 &= (I - KH)FP_1F^T(I - KH)^T + (I - KH)Q(I - KH)^T \\
P_2 &= (I - KH)FP_2F^T(I - KH)^T + KRK^T
\end{aligned}
$$

- Adding the two together and it shows linearity in predictor/corrector type filter.

$$
\begin{aligned}
P_1 + P_2 &= (I - KH)FP_1F^T(I - KH)^T + (I - KH)Q(I - KH)^T + \\
&\quad (I - KH)FP_2F^T(I - KH)^T + KRK^T \\
&= (I - KH)F(P_1 + P_2)F^T(I - KH)^T + \\
&\quad (I - KH)Q(I - KH)^T + KRK^T \qquad = P
\end{aligned}
$$

- Assume the true noise covariance are different from Q and R. Let

$$
\begin{aligned}
\tilde{Q} &= (1 + \alpha)Q \\
\tilde{R} &= (1 + \beta)R
\end{aligned}
$$

Robust Kalman Filter

- P should be updated with a ΔP.

$$\begin{aligned}\tilde{P} &= (I-KH)F\tilde{P}F^T(I-KH)^T + (I-KH)\tilde{Q}(I-KH)^T + K\tilde{R}K^T \\ P+\Delta P &= (I-KH)F(P+\Delta P)F^T(I-KH)^T + \\ &\quad (1+\alpha)(I-KH)Q(I-KH)^T + (1+\beta)KRK^T\end{aligned}$$

- Comparing to the original equation:

$$\Delta P = (I-KH)F\Delta PF^T(I-KH)^T + \alpha(I-KH)Q(I-KH)^T + \beta KRK^T$$

- Similarly, we assume P_1 and P_2 are modified with ΔP_1 and ΔP_2.

$$\begin{aligned}\tilde{P}_1 &= (I-KH)F\tilde{P}_1F^T(I-KH)^T + (I-KH)\tilde{Q}(I-KH)^T \\ P_1+\Delta P_1 &= (I-KH)F(P_1+\Delta P_1)F^T(I-KH)^T + \\ &\quad (1+\alpha)(I-KH)Q(I-KH)^T \\ \tilde{P}_2 &= (I-KH)F\tilde{P}_2F^T(I-KH)^T + K\tilde{R}K^T \\ P_2+\Delta P_2 &= (I-KH)F(P_2+\Delta P_2)F^T(I-KH)^T + (1+\beta)KRK^T \\ \Delta P_1 &= (I-KH)F\Delta P_1F^T(I-KH)^T + \alpha(I-KH)Q(I-KH)^T \\ \Delta P_2 &= (I-KH)F\Delta P_2F^T(I-KH)^T + \beta KRK^T\end{aligned}$$

* $\Delta P = \Delta P_1 + \Delta P_2$

$$\boxed{\Delta P = \alpha P_1 + \beta P_2}$$

Robust Kalman Filter

- Note that α and β are independent zero-mean R. V. with variance σ_1^2 and σ_2^2.

$$\begin{aligned} E[\mathrm{Tr}(\Delta P)] &= E(\alpha)\mathrm{Tr}(P_1) + E(\beta)\mathrm{Tr}(P_2) \\ &= 0 \\ E\left\{[\mathrm{Tr}(\Delta P)]^2\right\} &= E\left\{[\alpha\mathrm{Tr}(X_1) + \beta\mathrm{Tr}(X_2)]^2\right\} \\ &= \sigma_1^2\mathrm{Tr}^2(P_1) + \sigma_2^2\mathrm{Tr}^2(P_2) \end{aligned}$$

- We can write the performance index as a sum of Tr(P) and $E[\mathrm{Tr}^2(\Delta P)]$,

$$\begin{aligned} J &= \rho\mathrm{Tr}(P) + (1-\rho)E\left\{[\mathrm{Tr}(\Delta P)]^2\right\} \\ &= \rho[\mathrm{Tr}(P_1) + \mathrm{Tr}(P_2)] + (1-\rho)\left[\sigma_1^2\mathrm{Tr}^2(P_1) + \sigma_2^2\mathrm{Tr}^2(P_2)\right] \end{aligned}$$

- where ρ is the relative importance.

Synchronization Error

- Consider a discrete time system:

$$
\begin{aligned}
x(k) &= F(k,k-1)x(k-1) + w(k,k-1) \\
y(k) &= H(k)x(k) + v(k)
\end{aligned}
$$

where F(k, k-1) is the transition matrix from time t= k-1 to k, similarly to w(k,k-1).

- The state space equation can be generalized as follows.

$$x(k) = F(k,k_0)x(k_0) + w(k,k_0) \longrightarrow x(k_0) = F(k_0,k)[x(k) - w(k,k_0)]$$

where k_0 is any time before k and $F(k_0,k) = F^{-1}(k,k_0)$.

- The noise covariance has catered for all process noise. $w(k,k_0) \sim [0, Q(k,k_0)]$
- At time k, we have the posteriori covariance of the estimate as follows.

$$
\begin{aligned}
\hat{x}(k) &= E[x(k)|y(1),\cdots,y(k)] \\
&= E[x(k)|Y(k)] \\
P(k) &= E\left\{[x(k)-\hat{x}(k)][x(k)-\hat{x}(k)]^T|Y(k)\right\}
\end{aligned}
$$

Synchronization Error

- Suppose an out-of-sequence measurement has arrived. The modified state estimate and its covariance are shown as follows.

$$\begin{aligned}\hat{x}(k|k_0) &= E[x(k)|Y(k), y(k_0)] \\ P(k|k_0) &= E\left\{[x(k) - \hat{x}(k, k_0)][x(k) - \hat{x}(k, k_0)]^T |Y(k), y(k_0)\right\}\end{aligned}$$

- The updated state estimates and its covariance with new measurement:

$$\begin{aligned}E[x(k_0)|Y(k)] &= F(k_0, k)E[x(k) - w(k, k_0)|Y(k)] \\ &= F(k_0, k)[\hat{x}^-(k) - \hat{w}(k, k_0)]\end{aligned}$$

  where $\hat{w}(k, k_0)$ is the expected value of cumulated noise from time k_0 to k conditioning on all measurement before k.

- Define a vector:

$$z(k) = \begin{bmatrix} x(k) \\ w(k, k_0) \end{bmatrix}$$

- Recall:

$$\begin{aligned}\mathrm{Cov}(a|c) &= E[(a - \bar{a})(a - \bar{a})^T |c] \\ \mathrm{Cov}(a, b|c) &= E[(a - \bar{a})(b - \bar{b})^T |c]\end{aligned}$$

Synchronization Error

- It is found that the mean and covariance of state propagates over time with the following KF structure.

$$
\begin{aligned}
\hat{x}^-(k) &= F(k-1)\hat{x}^+(k-1) \\
P^-(k) &= F(k-1)P^+(k-1)F^T(k-1) + Q(k) \\
P_{xy} &= P^-(k)H^T(k) \\
P_y &= H(k)P^-(k)H^T(k) + R(k) \\
K(k) &= P_{xy}P_y^{-1} \\
\hat{x}^+(k) &= \hat{x}^-(k) + K(k)(y(k) - H(k)\hat{x}^-(k)) \\
P^+(k) &= P^-(k) - K(k)P_yK^T(k) \\
&= P^-(k) - P_{xy}P_y^{-1}P_{xy}^T
\end{aligned}
$$

- Generalize the above equations,

$$
\begin{aligned}
\hat{z}(k) &= \hat{z}^-(k) + \\
&\mathrm{Cov}[z(k), y(k)|Y(k-1)]\mathrm{Cov}^{-1}[y(k)|Y(k-1)]\left(y(k) - H(k)\hat{x}^-(k)\right) \\
\mathrm{Cov}[z(k)|Y(k)] &= \mathrm{Cov}[z(k)|Y(k-1)] - \\
&\mathrm{Cov}[z(k), y(k)|Y(k-1)]\mathrm{Cov}^{-1}[y(k)|Y(k-1)]\mathrm{Cov}[y(k), z(k)|Y(k-1)]
\end{aligned}
$$

Synchronization Error

- The covariance matrix can be written as follows.

$$\mathrm{Cov}[z(k), y(k)|Y(k-1)] = \begin{bmatrix} \mathrm{Cov}[x(k), y(k)|Y(k-1)] \\ \mathrm{Cov}[w(k,k_0), y(k)|Y(k-1)] \end{bmatrix}$$

$$\begin{aligned}
\mathrm{Cov}[x(k), y(k)|Y(k-1)] &= \mathrm{Cov}\left\{x(k)(H(k)x(k)+v(k)]^T|Y(k-1)\right\} \\
&= \mathrm{Cov}\left\{x(k)[H(k)x(k)]^T|Y(k-1)\right\} \\
&= \mathrm{Cov}\left\{x(k)\right\} H^T(k) \\
&= P^-(k)H^T(k)
\end{aligned}$$

$$\begin{aligned}
&\mathrm{Cov}[w(k,k_0), y(k)|Y(k-1)] = E\left\{w(k,k_0)[y(k)-\hat{y}^-(k)]^T|Y(k-1)\right\} \\
&= E\left\{w(k,k_0)[H(k)(F(k,k_0)x(k_0)+w(k,k_0))+v(k)-\hat{y}^-(k)]^T|Y(k-1)\right\} \\
&= E\left\{w(k,k_0)w^T(k,k_0)H^T(k)\right\} \\
&= Q(k,k_0)H^T(k)
\end{aligned}$$

- where the cross covariance of w(k, k_0) and x(k), v(k) and $\hat{y}_k^-$ = 0 with independence.

Synchronization Error

- Let $\hat{y}^{-}(k)$ be the expectation of y from time t = 0 to k-1.
- It is shown that $\text{Cov}[y(k)|Y(k-1)] = H(k)P^{-}(k)H^{T}(k) + R(k)$

$$\text{Cov}[r(k)] = S(k) \qquad r(k) = y(k) - H(k)\hat{x}^{-}(k)$$

$$\begin{aligned} \hat{z}(k) &= \begin{bmatrix} \hat{x}(k) \\ \hat{w}(k,k_0) \end{bmatrix} \\ &= \begin{bmatrix} \hat{x}^{-}(k) \\ \hat{w}^{-}(k,k_0) \end{bmatrix} + \begin{bmatrix} P^{-}(k)H^{T}(k) \\ Q(k,k_0)H^{T}(k) \end{bmatrix} S^{-1}(k)r(k) \end{aligned}$$

$$\begin{aligned} \hat{w}(k,k_0) &= \hat{w}^{-}(k,k_0) + Q(k,k_0)H^{T}(k)S^{-1}(k)r(k) \\ &= Q(k,k_0)H^{T}(k)S^{-1}(k)r(k) \end{aligned}$$

- Note that $E[\hat{w}(k,k_0)|Y(k-1)] = 0$

$$E[x(k_0)|Y(k)] = F(k_0,k)\left[\hat{x}(k) - Q(k,k_0)H^{T}(k)S^{-1}(k)r(k)\right]$$

- This is retrodiction of state estimate from time k back to k_0, i.e. use future state to guess back historical states.

Synchronization Error

- Recall:

$$\mathrm{Cov}[z(k)|Y(k)] = \mathrm{Cov}[z(k)|Y(k-1)] -$$
$$\mathrm{Cov}[z(k), y(k)|Y(k-1)]\mathrm{Cov}^{-1}[y(k)|Y(k-1)]\mathrm{Cov}[y(k), z(k)|Y(k-1)]$$

$$\begin{aligned}
\mathrm{Cov}[z(k)|Y(k)] &= \mathrm{Cov}\left\{\begin{bmatrix} x(k) \\ w(k,k_0) \end{bmatrix} |Y(k)\right\} \\
&= \mathrm{Cov}\left\{\begin{bmatrix} x(k) \\ w(k,k_0) \end{bmatrix} |Y(k-1)\right\} - \\
&\quad \mathrm{Cov}\left\{\begin{bmatrix} x(k) \\ w(k,k_0) \end{bmatrix}, y(k)|Y(k-1)\right\} \mathrm{Cov}^{-1}[y(k)|Y(k-1)] \times \\
&\quad \mathrm{Cov}\left\{y(k), \begin{bmatrix} x(k) \\ w(k,k_0) \end{bmatrix}^T |Y(k-1)\right\}
\end{aligned}$$

$$= \begin{bmatrix} \mathrm{Cov}[x(k)|Y(k-1)] & \mathrm{Cov}[x(k), w(k,k_0)|Y(k-1)] \\ \mathrm{Cov}[w(k,k_0), x(k)|Y(k-1)] & \mathrm{Cov}[w(k,k_0)|Y(k-1)] \end{bmatrix} -$$
$$\begin{bmatrix} \mathrm{Cov}[x(k), y(k)|Y(k-1)] \\ \mathrm{Cov}[w(k,k_0), y(k)|Y(k-1)] \end{bmatrix} \mathrm{Cov}^{-1}[y(k)|Y(k-1)] \times$$
$$\begin{bmatrix} \mathrm{Cov}[x(k), y(k)|Y(k-1)] \\ \mathrm{Cov}[w(k,k_0), y(k)|Y(k-1)] \end{bmatrix}^T$$

Synchronization Error

- Recall: $$x(k) = F(k, k_0)x(k_0) + w(k, k_0)$$

$$\begin{aligned}
\mathrm{Cov}[x(k), w(k, k_0)|Y(k-1)] &= E[x(k)w^T(k, k_0)|Y(k-1)] \\
&= E\left\{[F(k, k_0)x(k_0) + w(k, k_0)]w^T(k, k_0)|Y(k-1)\right\} \\
&= E\left\{w(k, k_0)w^T(k, k_0)|Y(k-1)\right\} \\
&= Q(k, k_0)
\end{aligned}$$

- Recall: $$\mathrm{Cov}[x(k), y(k)|Y(k-1)] = P^-(k)H^T(k)$$

$$\mathrm{Cov}[w(k, k_0), y(k)|Y(k-1)] = Q(k, k_0)H^T(k)$$

$$\mathrm{Cov}[y(k)|Y(k-1)] = H(k)P^-(k)H^T(k) + R(k)$$

$$\begin{aligned}
\mathrm{Cov}\left\{\begin{bmatrix} x(k) \\ w(k, k_0) \end{bmatrix} |Y(k)\right\} &= \begin{bmatrix} P^-(k) & Q(k, k_0) \\ Q(k, k_0) & Q(k, k_0) \end{bmatrix} - \\
&\begin{bmatrix} P^-(k)H^T(k) \\ Q(k, k_0)H^T(k) \end{bmatrix} S^{-1}(k) \begin{bmatrix} P^-(k)H^T(k) \\ Q(k, k_0)H^T(k) \end{bmatrix}^T
\end{aligned}$$

Synchronization Error

- The conditional covariance of $w(k,k_0)$ and cross covariance of $x(k)$ and $w(k, k_0)$ are

$$
\begin{aligned}
P_w(k,k_0) &= \mathrm{Cov}[w(k,k_0)|Y(k)] \\
&= Q(k,k_0) - Q(k,k_0)H^T(k)S^{-1}(k)H(k)Q(k,k_0) \\
P_{xw}(k,k_0) &= \mathrm{Cov}[x(k), w(k,k_0)|Y(k)] \\
&= Q(k,k_0) - P^-(k)H^T(k)S^{-1}(k)H(k)Q(k,k_0)
\end{aligned}
$$

- $x(k_0) = F(k_0,k)[x(k) - w(k,k_0)]$ gives the retrodiction as follows.

$$
\begin{aligned}
P(k_0,k) &= \mathrm{Cov}[x(k_0)|Y(k)] \\
&= F(k_0,k)\mathrm{Cov}[x(k) - w(k,k_0)|Y(k)]F^T(k_0,k) \\
&= F(k_0,k)\big\{\mathrm{Cov}[x(k)|Y(k)] - \mathrm{Cov}[x(k), w(k,k_0)|Y(k)] - \\
&\quad \mathrm{Cov}^T[x(k), w(k,k_0)|Y(k)] + \mathrm{Cov}[w(k,k_0)|Y(k)]\big\}F^T(k_0,k) \\
&= F(k_0,k)\left\{P^+(k) - P_{xw}(k,k_0) - P_{xw}^T(k,k_0) + \right. \\
&\quad \left. P_w(k,k_0)\right\}F^T(k_0,k)
\end{aligned}
$$

Synchronization Error

- The system equation $\begin{aligned} x(k) &= F(k,k-1)x(k-1) + w(k,k-1) \\ y(k) &= H(k)x(k) + v(k) \end{aligned}$ gives

$$\begin{aligned} S(k_0) &= \mathrm{Cov}[y(k_0)|Y(k)] \\ &= E\left\{[H(k_0)x(k_0) + v(k_0)][H(k_0)x(k_0) + v(k_0)]^T|Y(k)\right\} \\ &= H(k_0)P(k_0,k)H^T(k_0) + R(k_0) \end{aligned}$$

- The conditional covariance between $x(k)$ and $y(k_0)$ is

$$\begin{aligned} P_{xy}(k,k_0) &= \mathrm{Cov}[x(k), y(k_0)|Y(k)] \\ &= \mathrm{Cov}\left\{x(k), H(k_0)F(k_0,k)[x(k) - w(k,k_0)] + v(k_0)|Y(k)\right\} \\ &= [P^+(k) - P_{xw}(k,k_0)]F^T(k_0,k)H^T(k_0) \end{aligned}$$

- Substitute into the top partition of $\hat{z}(k)$ and $\mathrm{Cov}[z(k)|Y(k)]$

$$\hat{x}(k,k_0) = \hat{x}(k) + P_{xy}(k,k_0)S^{-1}(k_0)[y(k_0) - H(k_0)\hat{x}(k_0,k)]$$

$$\begin{aligned} \mathrm{Cov}[x(k)|Y(k), y(k_0)] &= P(k,k_0) \\ &= P(k) - P_{xy}(k,k_0)S^{-1}(k_0)P_{xy}^T(k,k_0) \end{aligned}$$

Synchronization Error

1. The Kalman filter is run normally on the basis of measurements that arrive sequentially. If we are presently at time k in the Kalman filter, then we have $\hat{x}^-(k)$ and $P^-(k)$, the *a priori* state estimate and covariance that are based on measurements up to and including time $(k-1)$. We also have $\hat{x}(k)$ and $P(k)$, the *a posteriori* state estimate and covariance that are based on measurements up to and including time k.

2. If we receive a measurement $y(k_0)$, where $k_0 < k$, then we can update the state estimate and its covariance to $\hat{x}(k, k_0)$ and $P(k, k_0)$ as follows.

   (a) Retrodict the state estimate from k back to k_0 as shown in Equation (10.118):

$$\begin{aligned} S(k) &= H(k)P^-(k)H^T(k) + R(k) \\ \hat{x}(k_0, k) &= F(k_0, k)\left[\hat{x}(k) - Q(k, k_0)H^T(k)S^{-1}(k)r(k)\right] \end{aligned} \quad (10.128)$$

   (b) Compute the covariance of the retrodicted state using Equations (10.122) and (10.123):

$$\begin{aligned} P_w(k, k_0) &= Q(k, k_0) - Q(k, k_0)H^T(k)S^{-1}(k)H(k)Q(k, k_0) \\ P_{xw}(k, k_0) &= Q(k, k_0) - P^-(k)H^T(k)S^{-1}(k)H(k)Q(k, k_0) \\ P(k_0, k) &= F(k_0, k)\left\{P(k) - P_{xw}(k, k_0) - P_{xw}^T(k, k_0) + \right. \\ &\quad \left. P_w(k, k_0)\right\} F^T(k_0, k) \end{aligned} \quad (10.129)$$

   (c) Compute the covariance of the retrodicted measurement at time k_0 using Equation (10.124):

$$S(k_0) = H(k_0)P(k_0, k)H^T(k_0) + R(k_0) \quad (10.130)$$

   (d) Compute the covariance of the state at time k and the retrodicted measurement at time k_0 using Equation (10.125):

$$P_{xy}(k, k_0) = [P(k) - P_{xw}(k, k_0)]F^T(k_0, k)H^T(k_0) \quad (10.131)$$

   (e) Use the delayed measurement $y(k_0)$ to update the state estimate and its covariance:

$$\begin{aligned} \hat{x}(k, k_0) &= \hat{x}(k) + P_{xy}(k, k_0)S^{-1}(k_0)[y(k_0) - H(k_0)\hat{x}(k_0, k)] \\ P(k, k_0) &= P(k) - P_{xy}(k, k_0)S^{-1}(k_0)P_{xy}^T(k, k_0) \end{aligned} \quad (10.132)$$

From Kalman Filter to H-inf Filter

The Kalman filter works well, but only under certain conditions.

- First, we need to know the mean and correlation of the noise w_k and v_k at each time instant.
- Second, we need to know the covariances Q_k and R_k of the noise processes. The Kalman filter uses Q_k and R_k as design parameters, so if we do not know Q_k and R_k then it may be difficult to successfully use a Kalman filter.
- Third, the attractiveness of the Kalman filter lies in the fact that it is the one estimator that results in the smallest possible standard deviation of the estimation error. That is, the Kalman filter is the minimum variance estimator if the noise is Gaussian, and it is the linear minimum variance estimator if the noise is not Gaussian. If we desire to minimize a different cost function (such as the worst-case estimation error) then the Kalman filter may not accomplish our objectives.
- Finally, we need to know the system model matrices F_k and H_k.

What if we don't have noise statistics?

What if we want to minimize the worst-case estimation, instead of the covariance of estimation error?

- H-inf filter is called minimax filter which minimizes the worst-case estimation error and it doesn't assume anything in noise statistics.

Constrained Optimization before H-inf Control

- Static Constrained Optimization:

$$\min_{x,w} J(x,w) \text{ such that } f(x,w)=0$$

Set the augmented cost function as $J_a = J + \lambda^T f$

It results in

$$\lambda^T = \frac{-\partial J}{\partial x}\left(\frac{\partial f}{\partial x}\right)^{-1}$$ (Lagrange Multiplier)

$$\frac{\partial J}{\partial w} - \frac{\partial J}{\partial x}\left(\frac{\partial f}{\partial x}\right)^{-1}\frac{\partial f}{\partial w} = 0$$

$$f = 0$$ f(x,w) = 0

$$\left.\frac{\partial J}{\partial x}\right|_{x^*,w^*}\Delta x + \left.\frac{\partial J}{\partial w}\right|_{x^*,w^*}\Delta w = 0$$

$$\left.\frac{\partial f}{\partial x}\right|_{x^*,w^*}\Delta x + \left.\frac{\partial f}{\partial w}\right|_{x^*,w^*}\Delta w = 0$$

$$\Delta x = -\left(\left.\frac{\partial f}{\partial x}\right|_{x^*,w^*}\right)^{-1}\left.\frac{\partial f}{\partial w}\right|_{x^*,w^*}\Delta w$$

e.g. $J(x,u) = x^2/2 + xu + u^2 + u$

$f(x,u) = x - 3 = 0$

$$\begin{aligned} J_a &= J + \lambda^T f \\ &= x^2/2 + xu + u^2 + u + \lambda(x-3) \end{aligned}$$

$$\begin{aligned} \frac{\partial J_a}{\partial x} &= x + u + \lambda = 0 \\ \frac{\partial J_a}{\partial u} &= x + 2u + 1 = 0 \\ \frac{\partial J_a}{\partial \lambda} &= x - 3 = 0 \end{aligned}$$

Constrained Optimization before H-inf Control

- Inequality Constraint: $\min J(x)$ such that $f(x) \leq 0$ =
  1. $\min J(x)$
  2. $\min J(x)$ such that $f(x) = 0$
- Dynamic Constrained Optimization:

$$x_{k+1} = F_k x_k + w_k \quad (k = 0, \cdots, N-1) \qquad J = \psi(x_0) + \sum_{k=0}^{N-1} \mathcal{L}_k(x, w)$$

Augmented Cost: $$J_a = \psi(x_0) + \sum_{k=0}^{N-1} \left[\mathcal{L}_k + \lambda_{k+1}^T (F_k x_k + w_k - x_{k+1})\right]$$

Hamiltonian: $$\mathcal{H}_k = \mathcal{L}_k + \lambda_{k+1}^T (F_k x_k + w_k)$$

$$\begin{aligned} \frac{\partial J_a}{\partial x_0} &= 0 \\ \frac{\partial J_a}{\partial x_N} &= 0 \end{aligned} \qquad \begin{aligned} \frac{\partial J_a}{\partial x_k} &= 0 \quad (k = 1, \cdots, N-1) \\ \frac{\partial J_a}{\partial w_k} &= 0 \quad (k = 0, \cdots, N-1) \\ \frac{\partial J_a}{\partial \lambda_k} &= 0 \quad (k = 0, \cdots, N) \end{aligned}$$

$$\longrightarrow$$

$$\boxed{\begin{aligned} \lambda_0^T + \frac{\partial \psi_0}{\partial x_0} &= 0 \\ -\lambda_N^T &= 0 \\ \lambda_k^T &= \frac{\partial \mathcal{H}_k}{\partial x_k} \quad (k = 1, \cdots, N-1) \\ \frac{\partial \mathcal{H}_k}{\partial w_k} &= 0 \quad (k = 0, \cdots, N-1) \end{aligned}}$$

H-inf Control with Game Theory Approach

- Let a linear discrete time system:

$$\begin{aligned} x_{k+1} &= F_k x_k + w_k \\ y_k &= H_k x_k + v_k \end{aligned}$$

- Our goal is to estimate a linear combination of states. $z_k = L_k x_k$
- We define the cost function as follows.

* For Kalman Filter, L_k = I.

$$J_1 = \frac{\sum_{k=0}^{N-1} ||z_k - \hat{z}_k||^2_{S_k}}{||x_0 - \hat{x}_0||^2_{P_0^{-1}} + \sum_{k=0}^{N-1} \left(||w_k||^2_{Q_k^{-1}} + ||v_k||^2_{R_k^{-1}} \right)}$$

- Our goal as engineer is to minimize J_1 with an estimate $\hat{z}_k$. Kalman Filter assumed nature is indifferent with pdf of noise given and obtain statistically optimal point.
- Nature's goal is to find disturbance w_k and v_k and initial state x_0 to maximize J_1. To avoid using infinity (i.e. brute force) to maximize J_1, the denominator is set as above. H-inf control assumes nature is preserved and actively degrades the state estimate.

H-inf Control with Game Theory Approach

- We try to find an estimate $\hat{z}_k$ which satisfies $J_1 < \frac{1}{\theta}$, where θ is user defined performance bound.
- The cost function can be written as

$$J = \frac{-1}{\theta}||x_0 - \hat{x}_0||^2_{P_0^{-1}} + \sum_{k=0}^{N-1}\left[||z_k - \hat{z}_k||^2_{S_k} - \frac{1}{\theta}\left(||w_k||^2_{Q_k^{-1}} + ||v_k||^2_{R_k^{-1}}\right)\right]$$
$$< 1$$

- The minimax equation is $J^* = \min_{\hat{z}_k} \max_{w_k, v_k, x_0} J$
- With $\hat{z}_k = L_k \hat{x}_k$, we choose $\hat{x}_k$ to minimize J. $\longrightarrow$ $J^* = \min_{\hat{x}_k} \max_{w_k, v_k, x_0} J$
- As y_0 depends on w_k, v_k and x_0, we can have $J^* = \min_{\hat{x}_k} \max_{w_k, y_k, x_0} J$
- Since $y_k = H_k x_k + v_k$, we see that $v_k = y_k - H_k x_k$ and

$$||v_k||^2_{R_k^{-1}} = ||y_k - H_k x_k||^2_{R_k^{-1}}$$

H-inf Control with Game Theory Approach

- Since $z_k = L_k x_k$ and $\hat{z}_k = L_k \hat{x}_k$, we see that

$$\begin{aligned} ||z_k - \hat{z}_k||^2_{S_k} &= (z_k - \hat{z}_k)^T S_k (z_k - \hat{z}_k) \\ &= (x_k - \hat{x}_k)^T L_k^T S_k L_k (x_k - \hat{x}_k) \\ &= ||x_k - \hat{x}_k||^2_{\bar{S}_k} \end{aligned}$$

  where $\bar{S}_k$ is defined as

$$\bar{S}_k = L_k^T S_k L_k$$

- The cost function can be written as:

$$\begin{aligned} J &= \frac{-1}{\theta} ||x_0 - \hat{x}_0||^2_{P_0^{-1}} + \sum_{k=0}^{N-1} \left[||x_k - \hat{x}_k||^2_{\bar{S}_k} - \frac{1}{\theta} \left(||w_k||^2_{Q_k^{-1}} + ||y_k - H_k x_k||^2_{R_k^{-1}} \right) \right] \\ &= \psi(x_0) + \sum_{k=0}^{N-1} \mathcal{L}_k \end{aligned}$$

H-inf Control with Game Theory Approach

- We first find a stationary point in J w.r.t. $\{x_0, w_k\}$, then a stationary point in J w.r.t. $\{\hat{x}_k, y_k\}$.
- Stationary with $\{x_0, w_k\}$: $J = \psi(x_0) + \sum_{k=0}^{N-1} \mathcal{L}_k$ subject to $x_{k+1} = F_k x_k + w_k$

Hamiltonian: $\mathcal{H}_k = \mathcal{L}_k + \frac{2\lambda_{k+1}^T}{\theta}(F_k x_k + w_k)$

$\frac{2\lambda_{k+1}^T}{\theta}$ is a time-varying Lagrange Multiplier.

Consider the constrained optimization problem,

$$
\begin{aligned}
\frac{2\lambda_0^T}{\theta} + \frac{\partial \psi_0}{\partial x_0} &= 0 \\
\frac{2\lambda_N^T}{\theta} &= 0 \\
\frac{\partial \mathcal{H}_k}{\partial w_k} &= 0 \\
\frac{2\lambda_k^T}{\theta} &= \frac{\partial \mathcal{H}_k}{\partial x_k}
\end{aligned}
$$

$$
\begin{aligned}
\frac{2\lambda_k}{\theta} &= 2\bar{S}_k(x_k - \hat{x}_k) + \frac{2}{\theta} H_k^T R_k^{-1}(y_k - H_k x_k) + \frac{2}{\theta} F_k^T \lambda_{k+1} \\
\lambda_k &= F_k^T \lambda_{k+1} + \theta \bar{S}_k (x_k - \hat{x}_k) + H_k^T R_k^{-1}(y_k - H_k x_k)
\end{aligned}
$$

$$
\begin{aligned}
\frac{2\lambda_0}{\theta} - \frac{2}{\theta} P_0^{-1}(x_0 - \hat{x}_0) &= 0 \\
P_0 \lambda_0 - x_0 + \hat{x}_0 &= 0 \\
x_0 &= \hat{x}_0 + P_0 \lambda_0
\end{aligned}
$$

$$\lambda_N = 0$$

$$
\begin{aligned}
-\frac{2}{\theta} Q_k^{-1} w_k + \frac{2}{\theta} \lambda_{k+1} &= 0 \\
w_k &= Q_k \lambda_{k+1}
\end{aligned}
$$

$$\boxed{x_{k+1} = F_k x_k + Q_k \lambda_{k+1}}$$

H-inf Control with Game Theory Approach

- From
$$\begin{aligned} \frac{2\lambda_0}{\theta} - \frac{2}{\theta}P_0^{-1}(x_0 - \hat{x}_0) &= 0 \\ P_0\lambda_0 - x_0 + \hat{x}_0 &= 0 \\ x_0 &= \hat{x}_0 + P_0\lambda_0 \end{aligned}$$
we know
$$\begin{aligned} x_0 &= \hat{x}_0 + P_0\lambda_0 \\ x_k &= \mu_k + P_k\lambda_k \end{aligned}.$$
μ_k and P_k are unknown.

- Initial Condition: $\mu_0 = \hat{x}_0.$

$$x_{k+1} = F_k x_k + Q_k\lambda_{k+1} \longrightarrow \mu_{k+1} + P_{k+1}\lambda_{k+1} = F_k\mu_k + F_kP_k\lambda_k + Q_k\lambda_{k+1}$$

$$x_k = \mu_k + P_k\lambda_k \longrightarrow \lambda_k = F_k^T\lambda_{k+1} + \theta\bar{S}_k(\mu_k + P_k\lambda_k - \hat{x}_k) + H_k^TR_k^{-1}[y_k - H_k(\mu_k + P_k\lambda_k)]$$

$$\begin{aligned} &\lambda_k - \theta\bar{S}_kP_k\lambda_k + H_k^TR_k^{-1}H_kP_k\lambda_k = \\ &\quad F_k^T\lambda_{k+1} + \theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k) \end{aligned}$$

$$\begin{aligned} \lambda_k \quad = \quad & \left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1} \times \\ & \left[F_k^T\lambda_{k+1} + \theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k)\right] \end{aligned}$$

$$\begin{aligned} \mu_{k+1} + P_{k+1}\lambda_{k+1} = F_k\mu_k + F_kP_k & \left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1} \times \\ & \left[F_k^T\lambda_{k+1} + \theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k)\right] + Q_k\lambda_{k+1} \end{aligned}$$

H-inf Control with Game Theory Approach

- Rearrange:

$$\mu_{k+1} - F_k\mu_k - F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1} \times$$
$$\left[\theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k)\right] =$$
$$\left[-P_{k+1} + F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1}F_k^T + Q_k\right]\lambda_{k+1}$$

- Setting LHS = RHS = 0,

$$\begin{aligned}\mu_{k+1} &= F_k\mu_k + F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1} \times \\ &\quad \left[\theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k)\right]\end{aligned}$$

$$\boxed{\mu_0 = \hat{x}_0}$$

$$\begin{aligned}P_{k+1} &= F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1}F_k^T + Q_k \\ &= F_k\tilde{P}_kF_k^T + Q_k\end{aligned}$$

$$\tilde{P}_k = \left[P_k^{-1} - \theta\bar{S}_k + H_k^TR_k^{-1}H_k\right]^{-1}$$

Solution:

$$\begin{aligned}x_0 &= \hat{x}_0 + P_0\lambda_0 \\ w_k &= Q_k\lambda_{k+1} \\ \lambda_N &= 0 \\ \lambda_k &= \left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1} \times \\ &\quad \left[F_k^T\lambda_{k+1} + \theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k)\right]\end{aligned}$$

$$\begin{aligned}P_{k+1} &= F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1}F_k^T + Q_k \\ \mu_0 &= \hat{x}_0 \\ \mu_{k+1} &= F_k\mu_k + F_kP_k\left[I - \theta\bar{S}_kP_k + H_k^TR_k^{-1}H_kP_k\right]^{-1} \times \\ &\quad \left[\theta\bar{S}_k(\mu_k - \hat{x}_k) + H_k^TR_k^{-1}(y_k - H_k\mu_k)\right]\end{aligned}$$

H-inf Control with Game Theory Approach

- Stationary w.r.t. $\hat{x}$ and y

minimize $J = \psi(x_k)|_{k=0} + \sum_{k=0}^{N-1} \mathcal{L}_k$

subject to $x_{k+1} = F_k x_k + w_k$

$$x_k = \mu_k + P_k \lambda_k$$
$$\mu_0 = \hat{x}_0$$

$$\begin{aligned} \lambda_k &= P_k^{-1}(x_k - \mu_k) \\ \lambda_0 &= P_0^{-1}(x_0 - \hat{x}_0) \end{aligned}$$

$$\begin{aligned} ||\lambda_0||^2_{P_0} &= \lambda_0^T P_0 \lambda_0 \\ &= (x_0 - \hat{x}_0)^T P_0^{-T} P_0 P_0^{-1}(x_0 - \hat{x}_0) \\ &= (x_0 - \hat{x}_0)^T P_0^{-1}(x_0 - \hat{x}_0) \\ &= ||x_0 - \hat{x}_0||^2_{P_0^{-1}} \end{aligned}$$

$$J = \frac{-1}{\theta}||\lambda_0||^2_{P_0} + \sum_{k=0}^{N-1}\left[||x_k - \hat{x}_k||^2_{\bar{S}_k} - \frac{1}{\theta}\left(||w_k||^2_{Q_k^{-1}} + ||y_k - H_k x_k||^2_{R_k^{-1}}\right)\right]$$

$$\boxed{x_k = \mu_k + P_k \lambda_k}$$

$$J = \frac{-1}{\theta}||\lambda_0||^2_{P_0} + \sum_{k=0}^{N-1}\left[||\mu_k + P_k\lambda_k - \hat{x}_k||^2_{\bar{S}_k} - \frac{1}{\theta}\left(\boxed{||w_k||^2_{Q_k^{-1}}} + ||y_k - H_k(\mu_k + P_k\lambda_k)||^2_{R_k^{-1}}\right)\right]$$

$$\begin{aligned} w_k^T Q_k^{-1} w_k &= \lambda_{k+1}^T Q_k^T Q_k^{-1} Q_k \lambda_{k+1} \\ &= \lambda_{k+1}^T Q_k \lambda_{k+1} \end{aligned}$$

H-inf Control with Game Theory Approach

- Note that Q_k is symmetric,

$$J = \frac{-1}{\theta}||\lambda_0||^2_{P_0} + \sum_{k=0}^{N-1}\left[||\mu_k + P_k\lambda_k - \hat{x}_k||^2_{\bar{S}_k} - \frac{1}{\theta}||y_k - H_k(\mu_k + P_k\lambda_k)||^2_{R_k^{-1}}\right] - \frac{1}{\theta}\sum_{k=0}^{N-1}||\lambda_{k+1}||^2_{Q_k}$$

- With $\lambda_N = 0$,

$$\sum_{k=0}^{N}\lambda_k^T P_k \lambda_k - \sum_{k=0}^{N-1}\lambda_k^T P_k \lambda_k = 0$$

$$\begin{aligned}
0 &= \lambda_0^T P_0 \lambda_0 + \sum_{k=1}^{N}\lambda_k^T P_k \lambda_k - \sum_{k=0}^{N-1}\lambda_k^T P_k \lambda_k \\
&= \lambda_0^T P_0 \lambda_0 + \sum_{k=0}^{N-1}\lambda_{k+1}^T P_{k+1} \lambda_{k+1} - \sum_{k=0}^{N-1}\lambda_k^T P_k \lambda_k \\
&= \frac{-1}{\theta}||\lambda_0||^2_{P_0} - \frac{1}{\theta}\sum_{k=0}^{N-1}(\lambda_{k+1}^T P_{k+1}\lambda_{k+1} - \lambda_k^T P_k \lambda_k)
\end{aligned}$$

H-inf Control with Game Theory Approach

- Subtract J with the 0 term,

$$
\begin{aligned}
J &= \sum_{k=0}^{N-1} \Big[||\mu_k + P_k\lambda_k - \hat{x}_k||^2_{\bar{S}_k} - \\
&\quad \frac{1}{\theta}||\lambda_{k+1}||^2_{Q_k} + \frac{1}{\theta}(\lambda_{k+1}^T P_{k+1}\lambda_{k+1} - \lambda_k^T P_k \lambda_k) - \frac{1}{\theta}||y_k - H_k(\mu_k + P_k\lambda_k)||^2_{R_k^{-1}} \Big] \\
&= \sum_{k=0}^{N-1} \Big[(\mu_k - \hat{x}_k)^T \bar{S}_k(\mu_k - \hat{x}_k) + 2(\mu_k - \hat{x}_k)^T \bar{S}_k P_k \lambda_k + \lambda_k^T P_k \bar{S}_k P_k \lambda_k + \\
&\quad \frac{1}{\theta}\boxed{\lambda_{k+1}^T (P_{k+1} - Q_k)\lambda_{k+1}} - \frac{1}{\theta}\lambda_k^T P_k \lambda_k - \frac{1}{\theta}(y_k - H_k\mu_k)^T R_k^{-1}(y_k - H_k\mu_k) + \\
&\quad \frac{2}{\theta}(y_k - H_k\mu_k)^T R_k^{-1} H_k P_k \lambda_k - \frac{1}{\theta}\lambda_k^T P_k H_k^T R_k^{-1} H_k P_k \lambda_k \Big]
\end{aligned}
$$

$$
\begin{aligned}
\lambda_{k+1}^T (P_{k+1} - Q_k)\lambda_{k+1} &= \lambda_{k+1}^T (Q_k + F_k \tilde{P}_k F_k^T - Q_k)\lambda_{k+1} \\
&= \lambda_{k+1}^T F_k \tilde{P}_k F_k^T \lambda_{k+1}
\end{aligned}
$$

$$\boxed{P_{k+1} = F_k \tilde{P}_k F_k^T + Q_k}$$

H-inf Control with Game Theory Approach

- Recall

$$\lambda_k = F_k^T \lambda_{k+1} + \theta \bar{S}_k (\mu_k + P_k \lambda_k - \hat{x}_k) + H_k^T R_k^{-1} [y_k - H_k(\mu_k + P_k \lambda_k)]$$

$$F_k^T \lambda_{k+1} = \lambda_k - \theta \bar{S}_k (\mu_k + P_k \lambda_k - \hat{x}_k) - H_k^T R_k^{-1} [y_k - H_k(\mu_k + P_k \lambda_k)]$$

- Substitute

$$(I - \theta P_k \bar{S}_k + P_k H_k^T R_k^{-1} H_k) = P_k \tilde{P}_k^{-1}$$

$$\begin{aligned}
&\lambda_{k+1}^T (P_{k+1} - Q_k) \lambda_{k+1} \\
&\quad = \left\{ \lambda_k^T P_k \tilde{P}_k^{-1} - \theta (\mu_k - \hat{x}_k)^T \bar{S}_k - (y_k - H_k \mu_k)^T R_k^{-1} H_k \right\} \\
&\qquad \tilde{P}_k \left\{ \lambda_k^T P_k \tilde{P}_k^{-1} - \theta (\mu_k - \hat{x}_k)^T \bar{S}_k - (y_k - H_k \mu_k)^T R_k^{-1} H_k \right\}^T \\
&\quad = \lambda_k^T P_k \tilde{P}_k^{-1} P_k \lambda_k - \theta (\mu_k - \hat{x}_k)^T \bar{S}_k P_k \lambda_k - (y_k - H_k \mu_k)^T R_k^{-1} H_k P_k \lambda_k - \\
&\qquad \theta \lambda_k P_k \bar{S}_k (\mu_k - \hat{x}_k) + \theta^2 (\mu_k - \hat{x}_k)^T \bar{S}_k \tilde{P}_k \bar{S}_k (\mu_k - \hat{x}_k) + \\
&\qquad \theta (y_k - H_k \mu_k)^T R_k^{-1} H_k \tilde{P}_k \bar{S}_k (\mu_k - \hat{x}_k) - \lambda_k^T P_k H_k^T R_k^{-1} (y_k - H_k \mu_k) + \\
&\qquad \theta (\mu_k - \hat{x}_k)^T \bar{S}_k \tilde{P}_k H_k^T R_k^{-1} (y_k - H_k \mu_k) + \\
&\qquad (y_k - H_k \mu_k)^T R_k^{-1} H_k \tilde{P}_k H_k^T R_k^{-1} (y_k - H_k \mu_k)
\end{aligned}$$

H-inf Control with Game Theory Approach

- Substitute: $\theta(\mu_k - \hat{x}_k)^T \bar{S}_k P_k \lambda_k = \theta \lambda_k^T P_k \bar{S}_k (\mu_k - \hat{x}_k)$ * P_k and S_k are symmetric.

$$\begin{aligned} \tilde{P}_k^{-1} &= \left[I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k\right] P_k^{-1} \\ &= P_k^{-1} \left[P_k^{-1} - \theta \bar{S}_k + H_k^T R_k^{-1} H_k\right] P_k^{-1} \\ &= P_k^{-1} \left[I - P_k \theta \bar{S}_k + P_k H_k^T R_k^{-1} H_k\right] \end{aligned}$$

$\rightarrow$

$$\begin{aligned} \lambda_k^T P_k \tilde{P}_k^{-1} P_k \lambda_k &= \lambda_k^T \left[I - \theta P_k \bar{S}_k + P_k H_k^T R_k^{-1} H_k\right] P_k \lambda_k \\ &= \lambda_k^T P_k \lambda_k - \theta \lambda_k^T P_k \bar{S}_k P_k \lambda_k + \lambda_k^T P_k H_k^T R_k^{-1} H_k P_k \lambda_k \end{aligned}$$

$$\begin{aligned} &\lambda_{k+1}^T (P_{k+1} - Q_k) \lambda_{k+1} \\ &\quad = \lambda_k^T P_k \lambda_k - \theta \lambda_k^T P_k \bar{S}_k P_k \lambda_k + \lambda_k^T P_k H_k^T R_k^{-1} H_k P_k \lambda_k - \\ &\qquad 2\theta (\mu_k - \hat{x}_k)^T \bar{S}_k P_k \lambda_k - 2(y_k - H_k \mu_k)^T R_k^{-1} H_k P_k \lambda_k + \\ &\qquad \theta^2 (\mu_k - \hat{x}_k)^T \bar{S}_k \tilde{P}_k \bar{S}_k (\mu_k - \hat{x}_k) + 2\theta (\mu_k - \hat{x}_k)^T \bar{S}_k \tilde{P}_k H_k^T R_k^{-1} (y_k - H_k \mu_k) + \\ &\qquad (y_k - H_k \mu_k)^T R_k^{-1} H_k \tilde{P}_k H_k^T R_k^{-1} (y_k - H_k \mu_k) \end{aligned}$$

$$\begin{aligned} J = \sum_{k=0}^{N-1} \Bigg[& (\mu_k - \hat{x}_k)^T \bar{S}_k (\mu_k - \hat{x}_k) - \frac{1}{\theta} (y_k - H_k \mu_k)^T R_k^{-1} (y_k - H_k \mu_k) + \\ & \theta (\mu_k - \hat{x})_k^T \bar{S}_k \tilde{P}_k \bar{S}_k (\mu_k - \hat{x}_k) + 2(\mu_k - \hat{x}_k)^T \bar{S}_k \tilde{P}_k H_k^T R_k^{-1} (y_k - H_k \mu_k) + \\ & \frac{1}{\theta} (y_k - H_k \mu_k)^T R_k^{-1} H_k \tilde{P}_k H_k^T R_k^{-1} (y_k - H_k \mu_k) \Bigg] \end{aligned}$$

H-inf Control with Game Theory Approach

- Hence, J = $\sum_{k=0}^{N-1}\Big[(\mu_k-\hat{x}_k)^T(\bar{S}_k+\theta\bar{S}_k\tilde{P}_k\bar{S}_k)(\mu_k-\hat{x}_k)+$
$2(\mu_k-\hat{x}_k)^T\bar{S}_k\tilde{P}_kH_k^TR_k^{-1}(y_k-H_k\mu_k)+$
$\frac{1}{\theta}(y_k-H_k\mu_k)^T(R_k^{-1}H_k\tilde{P}_kH_k^TR_k^{-1}-R_k^{-1})(y_k-H_k\mu_k)\Big]$

$$\begin{aligned}\frac{\partial J}{\partial \hat{x}_k} &= 2(\bar{S}_k+\theta\bar{S}_k\tilde{P}_k\bar{S}_k)(\hat{x}_k-\mu_k)+2\bar{S}_k\tilde{P}_kH_k^TR_k^{-1}(H_k\mu_k-y_k)\\ &= 0\\ \frac{\partial J}{\partial y_k} &= \frac{2}{\theta}(R_k^{-1}H_k\tilde{P}_kH_k^TR_k^{-1}-R_k^{-1})(y_k-H_k\mu_k)+2R_k^{-1}H_k\tilde{P}_k\bar{S}_k(\mu_k-\hat{x}_k)\\ &= 0\end{aligned}$$

$$\frac{\partial^2 J}{\partial \hat{x}_k^2} = 2(\bar{S}_k+\theta\bar{S}_k\tilde{P}_k\bar{S}_k)$$

Extreme Value (Min? or Max?)

$$\boxed{\begin{aligned}\hat{x}_k &= \mu_k\\ y_k &= H_k\mu_k\end{aligned}}$$

$\hat{x}_k^2$ is always positive definite.
→ minimum.

H-inf Control with Game Theory Approach

1. The system equations are given as

$$\begin{aligned} x_{k+1} &= F_k x_k + w_k \\ y_k &= H_k x_k + v_k \\ z_k &= L_k x_k \end{aligned} \tag{11.87}$$

where w_k and v_k are noise terms, and our goal is to estimate z_k.

2. The cost function is given as

$$J_1 = \frac{\sum_{k=0}^{N-1} \|z_k - \hat{z}_k\|^2_{S_k}}{\|x_0 - \hat{x}_0\|^2_{P_0^{-1}} + \sum_{k=0}^{N-1}\left(\|w_k\|^2_{Q_k^{-1}} + \|v_k\|^2_{R_k^{-1}}\right)} \tag{11.88}$$

where P_0, Q_k, R_k, and S_k are symmetric, positive definite matrices chosen by the engineer based on the specific problem.

3. The cost function can be made to be less than $1/\theta$ (a user-specified bound) with the following estimation strategy, which is derived from Equations (11.44), (11.60), (11.62), and (11.84):

$$\begin{aligned} \bar{S}_k &= L_k^T S_k L_k \\ K_k &= P_k \left[I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k\right]^{-1} H_k^T R_k^{-1} \\ \hat{x}_{k+1} &= F_k \hat{x}_k + F_k K_k (y_k - H_k \hat{x}_k) \\ P_{k+1} &= F_k P_k \left[I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k\right]^{-1} F_k^T + Q_k \end{aligned} \tag{11.89}$$

4. The following condition must hold at each time step ***k*** in order for the above estimator to be a solution to the problem:

$$P_k^{-1} - \theta \bar{S}_k + H_k^T R_k^{-1} H_k > 0$$

Steady-State H-inf Filter

- If the underlying system and design parameter is time-invariant, it is possible to obtain a steady-state solution to H-inf filtering problem.
- Consider a system:

$$
\begin{aligned}
x_{k+1} &= Fx_k + w_k \\
y_k &= Hx_k + v_k \\
z_k &= Lx_k
\end{aligned}
$$

- Our goal is to eliminate z_k such that $\lim_{N\to\infty} \frac{\sum_{k=0}^{N-1} ||z_k - \hat{z}_k||_S^2}{\sum_{k=0}^{N-1} \left(||w_k||_{Q^{-1}}^2 + ||v_k||_{R^{-1}}^2 \right)} < \frac{1}{\theta}$

- The cost function is modified as

$$
\begin{aligned}
\bar{S}_k &= L_k^T S_k L_k \\
K_k &= P_k \left[I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k\right]^{-1} H_k^T R_k^{-1} \\
\hat{x}_{k+1} &= F_k \hat{x}_k + F_k K_k (y_k - H_k \hat{x}_k) \\
P_{k+1} &= F_k P_k \left[I - \theta \bar{S}_k P_k + H_k^T R_k^{-1} H_k P_k\right]^{-1} F_k^T + Q_k
\end{aligned}
\quad \longrightarrow \quad
\begin{aligned}
\bar{S} &= L^T S L \\
K &= P \left[I - \theta \bar{S} P + H^T R^{-1} H P\right]^{-1} H^T R^{-1} \\
\hat{x}_{k+1} &= F \hat{x}_k + F K_k (y_k - H \hat{x}_k) \\
P &= F P \left[I - \theta \bar{S} P + H^T R^{-1} H P\right]^{-1} F^T + Q
\end{aligned}
$$

Steady-State H-inf Filter

- The condition must hold for the above estimator is a solution for the problem:

$$P^{-1} - \theta\bar{S} + H^T R^{-1} H > 0$$

- If θ, L, R or S is too large, or if H is too small, the H-inf estimator will have no solution.
- Note that P can be written as $P = F\left[P^{-1} - \theta\bar{S} + H^T R^{-1} H\right]^{-1} F^T + Q$
- The discrete-time algebraic Riccati Equation is

$$\begin{aligned} P &= F\left\{P - P\left[(H^T R^{-1} H - \theta\bar{S})^{-1} + P\right]^{-1} P\right\} F^T + Q \\ &= FPF^T - FP\left[(H^T R^{-1} H - \theta\bar{S})^{-1} + P\right]^{-1} PF^T + Q \end{aligned}$$

- Steady-State H-inf Filter will have large saving in computation effort.

Transfer Function Bound of H-inf Filter

- Consider a system with input u, transfer function G(z) and output x. If the sampling time is T and input signal u is comprised signals at frequency w such that the phase of u is defined as ϕ = Tw, the gain from u to x is defined as

$$\sup_{u \neq 0} \frac{||x||_2}{||u||_2} = \sigma_1 \left[G\left(e^{j\phi}\right) \right]$$

  where σ_1 is the largest singular value of the matrix G.

- If u can be comprised of arbitrary mix of frequencies, the max gain from u to x is

$$\begin{aligned} \sup_{\phi} \frac{||x||_2}{||u||_2} &= \sup_{\phi} \sigma_1 \left[G\left(e^{j\phi}\right) \right] \\ &= ||G||_\infty \quad \text{(infinity norm)} \end{aligned}$$

- Recall the steady-state H-inf filter, the cost function is $J = \lim_{N \to \infty} \frac{\sum_{k=0}^{N-1} ||z_k - \hat{z}_k||_S^2}{\sum_{k=0}^{N-1} \left(||w_k||_{Q^{-1}}^2 + ||v_k||_{R^{-1}}^2 \right)}$

Transfer Function Bound of H-inf Filter

- Since the H-inf filter makes the scalar smaller than $1/\theta$ for all w_k and v_k.

We can write

$$\|G_{\tilde{z}e}\|_\infty^2 = \sup_\phi \frac{\|z-\hat{z}\|_2^2}{\|w\|_2^2 + \|v\|_2^2} \le \frac{1}{\theta}$$

where

$\tilde{z} = z - \hat{z}$, $e^T = \begin{bmatrix} w^T & v^T \end{bmatrix}^T$

G_{ze} has input e and output z.

Example

Consider the system and filter discussed in Example 11.2:

$$\begin{aligned} x_{k+1} &= x_k + w_k \\ y_k &= x_k + v_k \\ \hat{x}_{k+1} &= (1-K)\hat{x}_k + K y_k \end{aligned}$$

The estimation error can be computed as

$$\begin{aligned} \tilde{x}_{k+1} &= x_{k+1} - \hat{x}_{k+1} \\ &= (1-K)\tilde{x}_k + w_k - K v_k \end{aligned}$$

Taking the z-transform of this equation gives

$$\begin{aligned} z\tilde{X}(z) &= (1-K)\tilde{X}(z) + W(z) - KV(z) \\ \tilde{X}(z) &= \frac{1}{z-1+K} \begin{bmatrix} 1 & -K \end{bmatrix} \begin{bmatrix} W(z) \\ V(z) \end{bmatrix} \\ &= G(z) \begin{bmatrix} W(z) \\ V(z) \end{bmatrix} \end{aligned}$$

Transfer Function Bound of H-inf Filter

Example (cont')

$G(z)$, the transfer function from w_k and v_k to $\tilde{x}_k$, is a 2×1 matrix. This matrix has one singular value, which is computed as

$$\begin{aligned}
\sigma^2(G) &= \lambda_{\max}\left[G(e^{j\phi})G^H(e^{j\phi})\right] \\
&= \frac{1+K^2}{(e^{j\phi}-1+K)(e^{-j\phi}-1+K)} \\
&= \frac{1+K^2}{K^2+2(K-1)(\cos\phi-1)} \qquad (11.113)
\end{aligned}$$

The supremum of this expression occurs at $\phi = 0$ when $K \leq 1$, so

$$\begin{aligned}
\|G\|_\infty^2 &= \sup_\phi \sigma^2\left[G\left(e^{j\phi}\right)\right] \\
&= \frac{1+K^2}{K^2} \qquad (11.114)
\end{aligned}$$

Therefore, if K = 1, θ = 1/2; K = 2/3, θ = 1

Note: When K increases, the infinity norm from the noise to the estimation error decreases. Yet, when K > 1, the estimator is unstable.

Continuous Time H-inf Filter

- Consider the system

$$\begin{aligned} \dot{x} &= Ax + Bu + w \\ y &= Cx + v \\ z &= Lx \end{aligned}$$

- In game theory approach, the cost function is

$$J_1 = \frac{\int_0^T ||z - \hat{z}||_S^2 \, dt}{||x(0) - \hat{x}(0)||_{P_0^{-1}}^2 + \int_0^T \left(||w||_{Q^{-1}}^2 + ||v||_{R^{-1}}^2 \right) dt} < \frac{1}{\theta}$$

- The estimator is

$$\begin{aligned} P(0) &= P_0 \\ \dot{P} &= AP + PA^T + Q - KCP + \theta P L^T S L P \\ K &= PC^T R^{-1} \\ \dot{\hat{x}} &= A\hat{x} + Bu + K(y - C\hat{x}) \\ \hat{z} &= L\hat{x} \end{aligned}$$

- The Riccati Equation is $\dot{P} = AP + PA^T + Q - PC^T R^{-1} CP + \theta P L^T S L P$

Transfer Function Approach

- Consider the LTI system:

$$\begin{aligned} x_{k+1} &= Fx_k + w_k \\ x_0 &= 0 \\ y_k &= Hx_k + v_k \\ z_k &= Lx_k \end{aligned}$$

$$\tilde{z}_k = z_k - \hat{z}_k$$

Augmented disturbance vector:

$$e_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix}$$

To find a steady state filter,
$$\begin{aligned} \|G_{\tilde{z}e}\|_\infty &= \sup_\omega \frac{\|\tilde{z}\|_2}{\|e\|_2} \\ &= \sup_\omega \sigma_1 \left[G_{\tilde{z}e}(j\omega)\right] \end{aligned}$$

The steady-state a priori to solve the problem is

$$\begin{aligned} P &= I + FPF^T - FPH^T(I + HPH^T)^{-1}HPF^T + \\ &\quad PL(I/\theta + LPL^T)^{-1}LP \\ K &= FPH^T(I + HPH^T)^{-1} \\ \hat{x}_{k+1} &= F\hat{x}_k + K(y_k - H\hat{x}_k) \end{aligned}$$

The steady-state a posteriori to solve is

$$\begin{aligned} \Sigma^{-1} &= \tilde{P}^{-1} - \theta L^T L + H^T H \\ \tilde{P} &= F\tilde{P}(H^T H\tilde{P} - \theta L^T L\tilde{P} + I)^{-1}F^T + I \\ \tilde{K} &= (I + \theta L^T L)^{-1}\Sigma H^T \\ &= \tilde{P}(I + H^T H\tilde{P})^{-1}H^T \\ \hat{x}_{k+1} &= F\hat{x}_k + \tilde{K}(y_{k+1} - HF\hat{x}_k) \end{aligned}$$

Transfer Function Approach

Note that P in a priori approach and $\tilde{P}$ in a posteriori approach are related as $P^{-1} = \tilde{P}^{-1} - \theta L^T L$

Define the 2n x 2n Hamiltonian as $\mathcal{H} = \begin{bmatrix} F^T + H^T H F^{-1} & \theta F^T L^T L - H^T H F^{-1}(I - \theta L^T L) \\ -F^{-1} & F^{-1}(I - \theta L^T L) \end{bmatrix}$

Denote the eigenvectors as ξ_i (i = 1, 2, …, n) and $\begin{bmatrix} \xi_1 & \cdots & \xi_n \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

$$P = X_2 X_1^{-1}$$

For a posteriori approach,

$$\tilde{\mathcal{H}} = \begin{bmatrix} F^{-T} & F^{-T}(H^T H - \theta L^T L) \\ F^{-T} & F + F^{-T}(H^T H - \theta L^T L) \end{bmatrix}$$

Denote the eigenvectors as $\tilde{\xi}_i$ (i = 1, 2, …, n) and $\begin{bmatrix} \tilde{\xi}_1 & \cdots & \tilde{\xi}_n \end{bmatrix} = \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix}$

$$\tilde{P} = \tilde{X}_2 \tilde{X}_1^{-1}$$

Kalman Filter VS H-inf Filter

- The advantages of H-inf estimation over Kalman filtering can be summarized as follows.
  1. H_∞ filtering provides a rigorous method for dealing with systems that have model uncertainty.
  2. H_∞ filtering provides a natural way to limit the frequency response of the estimator.
- The disadvantages of H, filtering compared to Kalman filtering can be summarized as follows.
  1. The filter performance is more sensitive to the design parameters.
  2. The theory underlying H_∞ filtering is more abstract and complicated.
- Applications:
  1. Systems in which stability margins must be guaranteed, or worst-case estimation performance is a primary consideration (rather than RMS estimation performance) [Sim96].
  2. Systems in which the model changes unpredictably, and identification and gain scheduling are too complex or time-consuming.
  3. Systems in which the model is not well known.

Mixed Kalman Filter/ H-inf Filter

- Recall the cost functions of Steady State KF and HF are

$$J_2 = \lim_{N\to\infty} \sum_{k=0}^{N} E\left(||x_k - \hat{x}_k||_2\right)$$

$$J_\infty = \lim_{N\to\infty} \max_{x_0, w_k, v_k} \frac{\sum_{k=0}^{N} ||x_k - \hat{x}_k||^2}{||x(0) - \hat{x}(0)||^2_{P_0^{-1}} + \sum_{k=0}^{N}\left(||w_k||^2_{Q_k^{-1}} + ||v_k||^2_{R_k^{-1}}\right)}$$

- Loosely speaking, KF minimizes RMS estimation error and HF minimizes worst-case estimation error.
- Consider an n-state observable LTI system:

$$\begin{aligned} x_{k+1} &= Fx_k + w_k \\ y_k &= Hx_k + v_k \end{aligned}$$

- Find an estimator $\hat{x}_{k+1} = \hat{F}x_k + Ky_k$ satisfying

1. $\hat{F}$ is a stable matrix (so the estimator is stable).
2. The H_∞ cost function is bounded by a user-specified parameter:

$$J_\infty < \frac{1}{\theta} \tag{12.5}$$

3. Among all estimators satisfying the above criteria, the filter minimizes the Kalman filter cost function J_2.

Mixed Kalman Filter/ H-inf Filter

1. Find the $n \times n$ positive semidefinite matrix P that satisfies the following Riccati equation:

$$P = FPF^T + Q + FP(I/\theta^2 - P)^{-1}PF^T - P_a V^{-1} P_a^T \tag{12.6}$$

where P_a and V are defined as

$$\begin{aligned} P_a &= FPH^T + FP(I/\theta^2 - P)^{-1}PH^T \\ V &= R + HPH^T + HP(I/\theta^2 - P)^{-1}PH^T \end{aligned} \tag{12.7}$$

2. Derive the $\hat{F}$ and K matrices in Equation (12.4) as

$$\begin{aligned} K &= P_a V^{-1} \\ \hat{F} &= F - KH \end{aligned} \tag{12.8}$$

3. The estimator $\hat{x}_{k+1} = \hat{F}x_k + Ky_k$ satisfies the mixed Kalman/H$_\infty$ estimation problem if and only if $\hat{F}$ is stable. In this case, the state estimation error satisfies the bound

$$\lim_{k\to\infty} E\left(\|x_k - \hat{x}_k\|^2\right) \le \mathrm{Tr}(P) \tag{12.9}$$

Robust Kalman/H-inf Filter

- Consider a system $\begin{aligned} x_{k+1} &= (F_k + \Delta F_k)x_k + w_k \\ y_k &= (H_k + \Delta H_k)x_k + v_k \end{aligned}$ with uncertainty $\begin{bmatrix} \Delta F_k \\ \Delta H_k \end{bmatrix} = \begin{bmatrix} M_{1k} \\ M_{2k} \end{bmatrix} \Gamma_k N_k$ which satisfies $\Gamma_k^T \Gamma_k \leq I$, ı.e. negatıve semıdefinite matrix.
- The state estimator should be in form of $\hat{x}_{k+1} = \hat{F}_k \hat{x}_k + K_k y_k$, with the following characteristics:

1. The estimator is stable (i.e., the eigenvalues of $\hat{F}_k$ are less than one in magnitude).

2. The estimation error $\tilde{x}_k$ satisfies the following worst-case bound:

$$\max_{w_k, v_k} \frac{\|\tilde{x}_k\|_2}{\|w_k\|_2 + \|v_k\|_2 + \|\tilde{x}_0\|_{S_1^{-1}} + \|x_0\|_{S_2^{-1}}} < \frac{1}{\theta} \tag{12.18}$$

3. The estimation error $\tilde{x}_k$ satisfies the following RMS bound:

$$E\left(\tilde{x}_k \tilde{x}_k^T\right) < P_k \tag{12.19}$$

Robust Kalman/H-inf Filter

1. Choose some scalar sequence $\alpha_k > 0$, and a small scalar $\epsilon > 0$.

2. Define the following matrices:

$$
\begin{aligned}
R_{11k} &= Q_k + \alpha_k M_{1k} M_{1k}^T \\
R_{12k} &= \alpha_k M_{1k} M_{2k}^T \\
R_{22k} &= R_k + \alpha_k M_{2k} M_{2k}^T
\end{aligned} \tag{12.20}
$$

3. Initialize P_k and $\tilde{P}_k$ as follows:

$$
\begin{aligned}
P_0 &= S_1 \\
\tilde{P}_0 &= S_2
\end{aligned} \tag{12.21}
$$

4. Find positive definite solutions P_k and $\tilde{P}_k$ satisfying the following Riccati equations:

$$
\begin{aligned}
P_{k+1} &= F_{1k} T_k F_{1k}^T + R_{11k} + R_{11k} R_{2k} R_{11k}^T - \\
&\quad \left[F_{1k} T_k H_{1k}^T + R_{11k} R_{2k} R_{12k} + R_{12k}\right] R_k^{-1} [\cdots]^T + \epsilon I \\
\tilde{P}_{k+1} &= F_k \tilde{P}_k F_k^T + F_k \tilde{P}_k N_k^T (\alpha_k I - N_k \tilde{P}_k N_k^T)^{-1} N_k \tilde{P}_k F_k^T + \\
&\quad R_{11k} + \epsilon I
\end{aligned} \tag{12.22}
$$

where the matrices R_{1k}, R_{2k}, F_{1k}, H_{1k}, and T_k are defined as

$$
\begin{aligned}
R_{1k} &= (\tilde{P}_k^{-1} - N_k^T N_k / \alpha_k)^{-1} F_k^T \\
R_{2k} &= R_{1k}^{-1} (\tilde{P}_k^{-1} - N_k^T N_k / \alpha_k)^{-1} R_{1k}^{-T} \\
F_{1k} &= F_k + R_{11k} R_{1k}^{-1} \\
H_{1k} &= H_k + R_{12k}^T R_{1k}^{-1} \\
T_k &= (P_k^{-1} - \theta^2 I)^{-1}
\end{aligned} \tag{12.23}
$$

5. If the Riccati equation solutions satisfy

$$
\begin{aligned}
\frac{1}{\theta^2} I &> P_k \\
\alpha_k I &> N_k \tilde{P}_k N_k^T
\end{aligned} \tag{12.24}
$$

then the estimator of Equation (12.17) solves the problem with

$$
\begin{aligned}
K_k &= \left[F_{1k} T_k H_{1k}^T + R_{11k} R_{2k} R_{12k} + R_{12k}\right] \tilde{R}_k^{-1} \\
\tilde{R}_k &= H_{1k} T_k H_{1k}^T + R_{12k}^T R_{2k} R_{12k} + R_{22k} \\
\hat{F}_k &= F_{1k} - K_k H_{1k}
\end{aligned} \tag{12.25}
$$

Constrained H-inf Filtering

- Consider a discrete LTI system: $\begin{aligned} x_{k+1} &= Fx_k + w_k + \delta_k \\ y_k &= Hx_k + v_k \end{aligned}$ subjected to $D_k x_k = d_k$
- Assume D_k is full rank and let $V_k = D_k^T D_k$
- The estimator is in form of predictor/corrector $\begin{aligned} \hat{x}_0 &= 0 \\ \hat{x}_{k+1} &= F\hat{x}_k + K_k(y_k - H\hat{x}_k) \end{aligned}$
- The noise δ_k is to maximize estimation error such that the input is

$$\delta_k = L_k[G_k(x_k - \hat{x}_k) + n_k]$$

- We assume the worst-case is $G_k = 0$ such that the input is all noise.
- The estimation error is defined as $e_k = x_k - \hat{x}_k$
- Hence

$$\begin{aligned} e_0 &= x_0 \\ e_{k+1} &= (F - K_k H + L_k G_k)e_k + w_k + L_k n_k - K_k v_k \end{aligned}$$

- With

$$D_k x_k = D_k \hat{x}_k = d_k$$

$$D_{k+1} F e_k = 0.$$

$$D_{k+1}^T D_{k+1} F e_k = V_{k+1} F e_k$$

$$\begin{aligned} e_0 &= x_0 \\ e_{k+1} &= [(I - V_{k+1})F - K_k H + L_k G_k]e_k + w_k + L_k n_k - K_k v_k \end{aligned}$$

* It is not acceptable as increasing e_k would increase L_k. Hence we decouple e.

Constrained H-inf Filtering

- Let $e_k = e_{1,k} + e_{2,k}$ such that $e_{1,k}$ and $e_{2,k}$ are evolved with

$$
\begin{aligned}
e_{1,0} &= x_0 \\
e_{1,k} &= [(I - V_{k+1})F - K_k H + L_k G k] e_{1,k} + w_k - K_k v_k \\
e_{2,0} &= 0 \\
e_{2,k} &= [(I - V_{k+1})F - K_k H + L_k G_k] e_{2,k} + L_k n_k
\end{aligned}
$$

- The objective function is defined as $J(K, L) = \mathrm{trace} \sum_{k=0}^{N} W_k E\left(e_{1,k} e_{1,k}^T - e_{2,k} e_{2,k}^T\right)$ where w_k is any positive weighting matrix.
- Again, it is aimed to find the optimal gain sequence K_k and L_k such that J(K, L) is minimized. The infinity norm is

$$\sup_{w_k, v_k} \frac{\sum_{k=0}^{N} e_k^T e_k}{\sum_{k=0}^{N} \left(w_k^T Q^{-1} w_k + v_k^T v_k\right)} < \frac{1}{\theta}$$

- The filtering solution should satisfy at the saddle point

$$J(K^*, L) \leq J(K^*, L^*) \leq J(K, L^*) \text{ for all } K, L$$

Constrained H-inf Filtering

- We define P_k and Σ_k as
$$\begin{aligned} P_0 &= E(x_0 x_0^T) \\ \Sigma_k &= (P_k H^T H - P_k G_k^T G_k + I)^{-1} P_k \\ P_{k+1} &= (I - V_{k+1}) F \Sigma_k F^T (I - V_{k+1}) + Q \end{aligned}$$
, which are non-singular.

- The following estimator satisfies the constrained H-inf filtering problem.
$$\begin{aligned} K_k^* &= (I - V_{k+1}) F \Sigma_k H^T \\ L_k^* &= (I - V_{k+1}) F \Sigma_k G_k^T \end{aligned}$$

- The estimator has a solution only when
$$(I - G_k P_k G_k^T) \geq 0$$

- The mean-square estimation error resulted in optimal gain K_k^* cannot be specified as it depends also on δ_k. Yet we can set an upper bound as game theory does.
$$E\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right] \leq P_k$$

- If it is unconstrained, set D and d = 0.
$$\begin{aligned} P_0 &= E(x_0 x_0^T) \\ P_k(I - H^T H \Sigma_k) &= (I - P_k G_k^T G_k)\Sigma_k \\ P_{k+1} &= F \Sigma_k F^T + Q \\ K_k^* &= F \Sigma_k C^T \\ L_k^* &= F \Sigma_k G_k^T \end{aligned}$$

Constrained H-inf Filtering

The constrained H_∞ filter can be summarized as follows.

1. We have a linear system given as

$$\begin{aligned} x_{k+1} &= F_k x_k + w_k \\ y_k &= H_k x_k + v_k \\ D_k x_k &= d_k \end{aligned} \tag{12.48}$$

   where w_k is the process noise, v_k is the measurement noise, and the last equation above specifies equality constraints on the state. We assume that the constraints are normalized so $D_k D_k^T = I$. The covariance of w_k is equal to Q_k, but w_k might have a zero mean or it might have a nonzero mean (i.e., it might contain a deterministic component). The covariance of v_k is the identity matrix.

2. Initialize the filter as follows:

$$\begin{aligned} \hat{x}_0 &= 0 \\ P_0 &= E(x_0 x_0^T) \end{aligned} \tag{12.49}$$

3. At each time step $k = 0, 1, \cdots$, do the following.

   (a) Choose the tuning parameter matrix G_k to weight the deterministic, biased component of the process noise. If $G_k = 0$ then we are assuming that the process noise is zero-mean and we get Kalman filter performance. As G_k increases we are assuming that there is more of a deterministic, biased component to the process noise. This gives us better worst-case error performance but worse RMS error performance.

   (b) Compute the next state estimate as follows:

$$\begin{aligned} V_k &= D_k^T D_k \\ \Sigma_k &= (P_k H_k^T H_k - P_k G_k^T G_k + I)^{-1} P_k \\ P_{k+1} &= (I - V_{k+1}) F_k \Sigma_k F_k^T (I - V_{k+1}) + Q_k \\ K_k &= (I - V_{k+1}) F_k \Sigma_k H_k^T \\ \hat{x}_{k+1} &= F_k \hat{x}_k + K_k (y_k - H_k \hat{x}_k) \end{aligned} \tag{12.50}$$

   (c) Verify that

$$(I - G_k P_k G_k^T) \geq 0 \tag{12.51}$$

   If not then the filter is invalid.

Extended Kalman Filter

- The derivation of linearized KF was based on linearizing the nonlinear system around a nominal state trajectory, which is an unknown.
- We linearize the nonlinear system around the Kalman filter estimate, and the Kalman filter estimate is based on the linearized system. This is the idea of the extended Kalman filter (EKF).
- Define nominal system trajectory

$$\begin{aligned} \dot{x}_0 &= f(x_0, u_0, w_0, t) \\ y_0 &= h(x_0, v_0, t) \end{aligned}$$

- The linearized KF equations:

$$\begin{aligned} \dot{x} &\approx f(x_0, u_0, w_0, t) + \left.\frac{\partial f}{\partial x}\right|_0 (x - x_0) + \left.\frac{\partial f}{\partial u}\right|_0 (u - u_0) + \\ &\quad \left.\frac{\partial f}{\partial w}\right|_0 (w - w_0) \\ &= f(x_0, u_0, w_0, t) + A\Delta x + B\Delta u + L\Delta w \\ y &\approx h(x_0, v_0, t) + \left.\frac{\partial h}{\partial x}\right|_0 (x - x_0) + \left.\frac{\partial h}{\partial v}\right|_0 (v - v_0) \\ &= h(x_0, v_0, t) + C\Delta x + M\Delta v \end{aligned}$$

$$\begin{aligned} \Delta\dot{x} &= A\Delta x + Lw \\ &= A\Delta x + \tilde{w} \\ \tilde{w} &\sim (0, \tilde{Q}), \qquad \tilde{Q} = LQL^T \\ \Delta y &= C\Delta x + Mv \\ &= C\Delta x + \tilde{v} \\ \tilde{v} &\sim (0, \tilde{R}), \qquad \tilde{R} = MRM^T \end{aligned}$$

$$\begin{aligned} \Delta\hat{x}(0) &= 0 \\ P(0) &= E\left[(\Delta x(0) - \Delta\hat{x}(0))(\Delta x(0) - \Delta\hat{x}(0))^T\right] \\ \Delta\dot{\hat{x}} &= A\Delta\hat{x} + K(\Delta y - C\Delta\hat{x}) \\ K &= PC^T\tilde{R}^{-1} \\ \dot{P} &= AP + PA^T + \tilde{Q} - PC^T\tilde{R}^{-1}CP \\ \hat{x} &= x_0 + \Delta\hat{x} \end{aligned}$$

Extended Kalman Filter

- From the linearized KF,

$$\dot{x}_0 + \Delta\dot{\hat{x}} = f(x_0, u_0, w_0, t) + A\Delta\hat{x} + K[y - y_0 - C(\hat{x} - x_0)]$$

- The nominal measurement is

$$\begin{aligned} y_0 &= h(x_0, v_0, t) \\ &= h(\hat{x}, v_0, t) \end{aligned} \longrightarrow \dot{\hat{x}} = f(\hat{x}, u, w_0, t) + K\left[y - h(\hat{x}, v_0, t)\right]$$

<u>Summary</u>

1. The system equations are given as

$$\begin{aligned} \dot{x} &= f(x, u, w, t) \\ y &= h(x, v, t) \\ w &\sim (0, Q) \\ v &\sim (0, R) \end{aligned} \tag{13.17}$$

Extended Kalman Filter

Summary

2. Compute the following partial derivative matrices evaluated at the current state estimate:

$$
\begin{aligned}
A &= \left.\frac{\partial f}{\partial x}\right|_{\hat{x}} \\
L &= \left.\frac{\partial f}{\partial w}\right|_{\hat{x}} \\
C &= \left.\frac{\partial h}{\partial x}\right|_{\hat{x}} \\
M &= \left.\frac{\partial h}{\partial v}\right|_{\hat{x}}
\end{aligned}
$$

3. Compute the following matrices:

$$
\begin{aligned}
\tilde{Q} &= LQL^T \\
\tilde{R} &= MRM^T
\end{aligned}
$$

4. Execute the following Kalman filter equations:

$$
\begin{aligned}
\hat{x}(0) &= E[x(0)] \\
P(0) &= E\left[(x(0) - \hat{x}(0))(x(0) - \hat{x}(0))^T\right] \\
\dot{\hat{x}} &= f(\hat{x}, u, w_0, t) + K\left[y - h(\hat{x}, v_0, t)\right] \\
K &= PC^T\tilde{R}^{-1} \\
\dot{P} &= AP + PA^T + \tilde{Q} - PC^T\tilde{R}^{-1}CP
\end{aligned}
$$

where the nominal noise values are given as $w_0 = 0$ and $v_0 = 0$.

Hybrid Extended Kalman Filter

- Many real engineering system is continuous in nature, but discrete in measurement.

The hybrid extended Kalman filter

1. The system equations with continuous-time dynamics and discrete-time measurements are given as follows:

$$\begin{aligned} \dot{x} &= f(x, u, w, t) \\ y_k &= h_k(x_k, v_k) \\ w(t) &\sim (0, Q) \\ v_k &\sim (0, R_k) \end{aligned} \tag{13.30}$$

2. Initialize the filter as follows:

$$\begin{aligned} \hat{x}_0^+ &= E[x_0] \\ P_0^+ &= E\left[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T\right] \end{aligned} \tag{13.31}$$

Hybrid Extended Kalman Filter

3. For $k = 1, 2, \cdots$, perform the following.

   (a) Integrate the state estimate and its covariance from time $(k-1)^+$ to time k^- as follows:

$$\begin{aligned} \dot{\hat{x}} &= f(\hat{x}, u, 0, t) \\ \dot{P} &= AP + PA^T + LQL^T \end{aligned} \qquad (13.32)$$

   where F and L are given in Equation (13.18). We begin this integration process with $\hat{x} = \hat{x}_{k-1}^+$ and $P = P_{k-1}^+$. At the end of this integration we have $\hat{x} = \hat{x}_k^-$ and $P = P_k^-$.

   (b) At time k, incorporate the measurement y_k into the state estimate and estimation covariance as follows:

$$\begin{aligned} K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + M_k R_k M_k^T)^{-1} \\ \hat{x}_k^+ &= \hat{x}_k^- + K_k(y_k - h_k(\hat{x}_k^-, 0, t_k)) \\ P_k^+ &= (I - K_k H_k) P_k^- (I - K_k H_k)^T + K_k M_k R_k M_k^T K_k^T \end{aligned} \qquad (13.33)$$

   H_k and M_k are the partial derivatives of $h_k(x_k, v_k)$ with respect to x_k and v_k, and are both evaluated at $\hat{x}_k^-$. Note that other equivalent expressions can be used for K_k and P_k^+, as is apparent from Equation (5.19).

$$\begin{aligned} A &= \left.\frac{\partial f}{\partial x}\right|_{\hat{x}} \\ L &= \left.\frac{\partial f}{\partial w}\right|_{\hat{x}} \\ C &= \left.\frac{\partial h}{\partial x}\right|_{\hat{x}} \\ M &= \left.\frac{\partial h}{\partial v}\right|_{\hat{x}} \end{aligned}$$

Discrete Time EKF

- Consider the system model

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k &= h_k(x_k, v_k) \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned}$$

with Taylor Series around

$$\begin{aligned} x_{k-1} &= \hat{x}_{k-1}^+ \\ w_{k-1} &= 0 \end{aligned}$$

$$\begin{aligned} x_k &= f_{k-1}(\hat{x}_{k-1}^+, u_{k-1}, 0) + \left.\frac{\partial f_{k-1}}{\partial x}\right|_{\hat{x}_{k-1}^+} (x_{k-1} - \hat{x}_{k-1}^+) + \left.\frac{\partial f_{k-1}}{\partial w}\right|_{\hat{x}_{k-1}^+} w_{k-1} \\ &= f_{k-1}(\hat{x}_{k-1}^+, u_{k-1}, 0) + F_{k-1}(x_{k-1} - \hat{x}_{k-1}^+) + L_{k-1}w_{k-1} \\ &= F_{k-1}x_{k-1} + \left[f_{k-1}(\hat{x}_{k-1}^+, u_{k-1}, 0) - F_{k-1}\hat{x}_{k-1}^+\right] + L_{k-1}w_{k-1} \\ &= F_{k-1}x_{k-1} + \tilde{u}_{k-1} + \tilde{w}_{k-1} \end{aligned}$$

- The known signal and noise signal are:

$$\begin{aligned} \tilde{u}_k &= f_k(\hat{x}_k^+, u_k, 0) - F_k\hat{x}_k^+ \\ \tilde{w}_k &\sim (0, L_k Q_k L_k^T) \end{aligned}$$

- Linearize the measurement equation:
- The known signal and noise signal are

$$\begin{aligned} y_k &= h_k(\hat{x}_k^-, 0) + \left.\frac{\partial h_k}{\partial x}\right|_{\hat{x}_k^-} (x_k - \hat{x}_k^-) + \left.\frac{\partial h_k}{\partial v}\right|_{\hat{x}_k^-} v_k \\ &= h_k(\hat{x}_k^-, 0) + H_k(x_k - \hat{x}_k^-) + M_k v_k \\ &= H_k x_k + \left[h_k(\hat{x}_k^-, 0) - H_k\hat{x}_k^-\right] + M_k v_k \\ &= H_k x_k + z_k + \tilde{v}_k \end{aligned}$$

$$\begin{aligned} z_k &= h_k(\hat{x}_k^-, 0) - H_k\hat{x}_k^- \\ \tilde{v}_k &\sim (0, M_k R_k M_k^T) \end{aligned}$$

Discrete Time EKF

1. The system and measurement equations are given as follows:

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k &= h_k(x_k, v_k) \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned}$$

2. Initialize the filter as follows:

$$\begin{aligned} \hat{x}_0^+ &= E(x_0) \\ P_0^+ &= E[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T] \end{aligned}$$

3. For $k = 1, 2, \cdots$, perform the following.

   (a) Compute the following partial derivative matrices:

$$\begin{aligned} F_{k-1} &= \left.\frac{\partial f_{k-1}}{\partial x}\right|_{\hat{x}_{k-1}^+} \\ L_{k-1} &= \left.\frac{\partial f_{k-1}}{\partial w}\right|_{\hat{x}_{k-1}^+} \end{aligned}$$

   (b) Perform the time update of the state estimate and estimation-error covariance as follows:

$$\begin{aligned} P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + L_{k-1}Q_{k-1}L_{k-1}^T \\ \hat{x}_k^- &= f_{k-1}(\hat{x}_{k-1}^+, u_{k-1}, 0) \end{aligned} \qquad (13.47)$$

   (c) Compute the following partial derivative matrices:

$$\begin{aligned} H_k &= \left.\frac{\partial h_k}{\partial x}\right|_{\hat{x}_k^-} \\ M_k &= \left.\frac{\partial h_k}{\partial v}\right|_{\hat{x}_k^-} \end{aligned} \qquad (13.48)$$

   (d) Perform the measurement update of the state estimate and estimation-error covariance as follows:

$$\begin{aligned} K_k &= P_k^-H_k^T(H_kP_k^-H_k^T + M_kR_kM_k^T)^{-1} \\ \hat{x}_k^+ &= \hat{x}_k^- + K_k[y_k - h_k(\hat{x}_k^-, 0)] \\ P_k^+ &= (I - K_kH_k)P_k^- \end{aligned} \qquad (13.49)$$

Iterated EKF

1. The nonlinear system and measurement equations are given as follows:

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k &= h_k(x_k, v_k) \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned}$$

2. Initialize the filter as follows.

$$\begin{aligned} \hat{x}_0^+ &= E(x_0) \\ P_0^+ &= E\left[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T\right] \end{aligned} \tag{13.60}$$

3. For $k = 1, 2, \cdots$, do the following.

   (a) Perform the following time-update equations:

$$\begin{aligned} P_k^- &= F_{k-1}P_{k-1}^+F_{k-1}^T + L_{k-1}Q_{k-1}L_{k-1}^T \\ \hat{x}_k^- &= f_{k-1}(\hat{x}_{k-1}^+, u_{k-1}, 0) \end{aligned} \tag{13.61}$$

   where the partial derivative matrices F_{k-1} and L_{k-1} are defined as follows:

$$\begin{aligned} F_{k-1} &= \left.\frac{\partial f_{k-1}}{\partial x}\right|_{\hat{x}_{k-1}^+} \\ L_{k-1} &= \left.\frac{\partial f_{k-1}}{\partial w}\right|_{\hat{x}_{k-1}^+} \end{aligned} \tag{13.62}$$

   Up to this point the iterated EKF is the same as the standard discrete-time EKF.

   (b) Perform the measurement update by initializing the iterated EKF estimate to the standard EKF estimate:

$$\begin{aligned} \hat{x}_{k,0}^+ &= \hat{x}_k^- \\ P_{k,0}^+ &= P_k^- \end{aligned} \tag{13.63}$$

   For $i = 0, 1, \cdots, N$, evaluate the following equations (where N is the desired number of measurement-update iterations):

$$\begin{aligned} H_{k,i} &= \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k,i}^+} \\ M_{k,i} &= \left.\frac{\partial h}{\partial v}\right|_{\hat{x}_{k,i}^+} \\ K_{k,i} &= P_k^- H_{k,i}^T \left(H_{k,i}P_k^- H_{k,i}^T + M_{k,i}R_k M_{k,i}^T\right)^{-1} \\ P_{k,i+1}^+ &= (I - K_{k,i}H_{k,i})P_k^- \\ \hat{x}_{k,i+1}^+ &= \hat{x}_k^- + K_{k,i}\left[y_k - h(\hat{x}_{k,i}^+) - H_{k,i}(\hat{x}_k^- - \hat{x}_{k,i}^+)\right] \end{aligned} \tag{13.64}$$

   (c) The final *a posteriori* state estimate and estimation-error covariance are given as follows:

$$\begin{aligned} \hat{x}_k^+ &= \hat{x}_{k,N+1}^+ \\ P_k^+ &= P_{k,N+1}^+ \end{aligned} \tag{13.65}$$

EKF2 – Second order approximation

1. The system equations are given as follows:

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, k) + w_k \\ y_k &= h(x_k, k) + v_k \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned}$$

2. The estimator is initialized as follows:

$$\begin{aligned} \hat{x}_0^+ &= E(x_0) \\ P_0^+ &= E\left[(x_0 - \hat{x}_0^+)(x_0 - \hat{x}_0^+)^T\right] \end{aligned}$$

3. The time update equations are given as follows:

$$\begin{aligned} \hat{x}_{k+1}^- &= f(\hat{x}_k^+, u_k, k) + \frac{1}{2}\sum_{i=1}^{n} \phi_i \mathrm{Tr}\left[\left.\frac{\partial^2 f_i}{\partial x}\right|_{\hat{x}_k^+} P_k^+\right] \\ P_{k+1}^- &= F P_k^+ F^T + Q_k \end{aligned}$$

$$\phi_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow i\text{th element} \\ \\ \\ \\ \end{array}$$

$$F = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_k^+}$$

4. The measurement update equations are given as follows:

$$\begin{aligned} \hat{x}_k^+ &= \hat{x}_k^- + K_k[y_k - h(\hat{x}_k^-, k)] - \pi_k \\ \pi_k &= \frac{1}{2} K_k \sum_{i=1}^{m} \phi_i \mathrm{Tr}\left[D_{k,i} P_k^-\right] \\ D_{k,i} &= \left.\frac{\partial^2 h_i(x_k, k)}{\partial x^2}\right|_{\hat{x}_k^-} \\ K_k &= P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k\right)^{-1} \\ H_k &= \left.\frac{\partial h(x_k, k)}{\partial x}\right|_{\hat{x}_k^-} \\ P_k^+ &= (I - K_k H_k) P_k^- \end{aligned}$$

Gaussian Sum Filter

- Approximate sum of non-Gaussian pdf with Gaussian.

1. The discrete-time n-state system and measurement equations are given as follows:

$$\begin{aligned} x_k &= f_{k-1}(x_{k-1}, u_{k-1}, w_{k-1}) \\ y_k &= h_k(x_k, v_k) \\ w_k &\sim (0, Q_k) \\ v_k &\sim (0, R_k) \end{aligned} \tag{13.102}$$

2. Initialize the filter by approximating the pdf of the initial state as follows:

$$\text{pdf}(\hat{x}_0^+) = \sum_{i=1}^{M} a_{0i} N(\hat{x}_{0i}^+, P_{0i}^+) \tag{13.103}$$

The a_{0i} coefficients (which are positive and add up to 1), the $\hat{x}_{0i}^+$ means, and the P_{0i}^+ covariances, are chosen by the user to provide a good approximation to the pdf of the initial state.

Gaussian Sum Filter

3 (a) The *a priori* state estimate is obtained by first executing the following time-update equations for $i = 1, \cdots, M$:

$$\begin{aligned}
\hat{x}_{ki}^- &= f_{k-1}(\hat{x}_{k-1,i}^+, u_{k-1}, 0) \\
F_{k-1,i} &= \left.\frac{\partial f_{k-1}}{\partial x_{k-1}}\right|_{\hat{x}_{k-1,i}^+} \\
P_{ki}^- &= F_{k-1,i} P_{k-1,i}^+ F_{k-1,i}^T + Q_{k-1} \\
a_{ki} &= a_{k-1,i}
\end{aligned} \tag{13.104}$$

The pdf of the *a priori* state estimate is obtained by the following sum:

$$\text{pdf}(\hat{x}_k^-) = \sum_{i=1}^{M} a_{ki} N(\hat{x}_{ki}^-, P_{ki}^-) \tag{13.105}$$

(b) The *a posteriori* state estimate is obtained by first executing the following measurement update equations for $i = 1, \cdots, M$:

$$\begin{aligned}
H_{ki} &= \left.\frac{\partial h_k}{\partial x_k}\right|_{\hat{x}_{ki}^-} \\
K_{ki} &= P_{ki}^- H_{ki}^T (H_{ki} P_{ki}^- H_{ki}^T + R_k)^{-1} \\
P_{ki}^+ &= P_{ki}^- - K_{ki} H_{ki} P_{ki}^- \\
\hat{x}_{ki}^+ &= \hat{x}_{ki}^- + K_{ki} \left[y_k - h_k(\hat{x}_{ki}^-, 0) \right]
\end{aligned} \tag{13.106}$$

The weighting coefficients a_{ki} for the individual estimates are obtained as follows:

$$\begin{aligned}
r_{ki} &= y_k - h_k(\hat{x}_{ki}^-, 0) \\
S_{ki} &= H_{ki} P_{ki}^- H_{ki}^T + R_k \\
\beta_{ki} &= \frac{\exp\left[-r_{ki}^T S_{ki}^{-1} r_{ki}/2 \right]}{(2\pi)^{n/2} |S_{ki}|^{1/2}} \\
a_{ki} &= \frac{a_{k-1,i}\beta_{ki}}{\sum_{j=1}^{M} a_{k-1,j}\beta_{kj}}
\end{aligned} \tag{13.107}$$

Note that the weighting coefficient a_{ki} is computed by using the measurement y_k to obtain the relative confidence β_{ki} of the estimate $\hat{x}_{ki}^-$. The pdf of the *a posteriori* state estimate is obtained by the following sum:

$$\text{pdf}(\hat{x}_k^+) = \sum_{i=1}^{M} a_{ki} N(\hat{x}_{ki}^+, P_{ki}^+) \tag{13.108}$$

Parameter Estimation

- Suppose that we have a discrete-time system model, but the system matrices depend in a nonlinear way on an unknown parameter vector ***p***

$$
\begin{aligned}
x_{k+1} &= F_k(p)x_k + G_k(p)u_k + L_k(p)w_k \\
y_k &= H_k x_k + v_k \\
p &= \text{ unknown parameter vector}
\end{aligned}
$$

- Use an augmented state vector x’ $x'_k = \begin{bmatrix} x_k \\ p_k \end{bmatrix}$
- If p_k is constant, we model $p_{k+1} = p_k + w_{pk}$ such that w_{pk} is artificial noise.
- Augmented System Model: $$
\begin{aligned}
x'_{k+1} &= \begin{bmatrix} F_k(p_k)x_k + G_k(p_k)u_k + L_k(p_k)w_k \\ p_k + w_{pk} \end{bmatrix} \\
&= f(x'_k, u_k, w_k, w_{pk}) \\
y_k &= \begin{bmatrix} H_k & 0 \end{bmatrix} \begin{bmatrix} x_k \\ p_k \end{bmatrix} + v_k
\end{aligned}
$$

- We can apply EKF to solve the problem.

Conclusion

- Kalman Filter with Riccati Equation, and its assumption
- Multiple State KF; Steady-State KF; Sequential KF; Information Filter; Square Root Update Equation; Colored Noise; Fading Memory Filter; Hamiltonian; Continuous KF; Wiener Filter; Reduced Order KF; Robust Filter
- Synchronization Error (i.e. Delay)
- H-inf Control with Game Theory; Constrained Optimization
- Steady State H-inf; Infinity Norm
- Mixed Kalman/H-inf Filter; Constrained H-inf Filtering
- Extended Kalman Filter; EKF2; Iterative Kalman Filter
- Parameter Estimation

Exercise

1. The measured output of a simple moving average process is $y_k = z_k + z_{k-1}$, where $\{z_j\}$ is zero-mean white noise with a variance of one.
   a) Generate a state-space description for this system with the first element of x_k equal to z_{k-1} and second element equal to z_k.
   b) Suppose that the initial estimation-error covariance is equal to the identity matrix. Show that the *a posteriori* estimation-error covariance is given by

$$P_k^+ = \frac{1}{k+1}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

   c) Find $E\left[\|x_k - \hat{x}_k^+\|_2^2\right]$ as a function of k.

2. Suppose that a Kalman filter is designed for the system

$$\begin{aligned} x_{k+1} &= x_k \\ y_k &= x_k + v_k \\ v_k &\sim (0, R) \end{aligned}$$

   a) Suppose that $E(x_0^2) = 1$. Design a Kalman filter for the system and find a closed-form expression for P_k^-. What is the limit of P_k^- as $k \to \infty$?
   b) Now suppose that the true process equation is actually $x_{k+1} = x_k + w_k$, where $w_k \sim (0, Q)$. Find a difference equation for the variance of the *a priori* estimation error if the Kalman filter that you designed in part (a) is used to estimate the state. What is the limit of the estimation-error variance as $k \to \infty$?

Exercise

3. This example is based on [Kam71]. Suppose that we have an LTI system with

$$
\begin{aligned}
P_k^- &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
H &= \begin{bmatrix} 1 & 0 \end{bmatrix} \\
F &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
Q &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}
\end{aligned} \tag{6.78}
$$

Find a priori and a posteriori Kalman Gain and Covariance at t = k and k+1.
Implement measurement update equation with Potter's Algorithm.

4. Suppose that you have a system with the following measurement and measurement noise covariance matrices:

$$
\begin{aligned}
H &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
R &= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}
\end{aligned}
$$

You want to use a sequential Kalman filter to estimate the state of the system. Derive the normalized measurement, measurement matrix, and measurement noise covariance matrix that could be used in a sequential Kalman filter.

Exercise

5. In this example we will use the continuous-time Kalman filter to estimate a constant given continuous-time noisy measurements:

$$\begin{aligned} \dot{x} &= 0 \\ y &= x + v \\ v &\sim (0, R) \end{aligned} \tag{8.38}$$

Find the steady state Kalman Gain, and covariance with state update equation.

6. Suppose that we want to estimate a gyroscope drift rate ϵ (assumed to be constant) given measurements of the gyro angle θ. The system and measurement model can be written as

$$\begin{aligned} \dot{\theta} &= \epsilon \\ y &= \theta + v \\ \begin{bmatrix} \dot{\theta} \\ \dot{\epsilon} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \epsilon \end{bmatrix} \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \epsilon \end{bmatrix} + v \\ v &\sim (0, R) \end{aligned} \tag{8.65}$$

Solve directly with Riccati Equation.

7. Find the steady-state solution of the differential Riccati equation for a scalar system. Show from your solution how the steady-state solution changes with A, C, Q, and R, and give intuitive explanations.

Exercise

8. Consider the system

$$
\begin{aligned}
x_k &= \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} x_{k-1} + w_{k-1} \\
y_k &= \begin{bmatrix} 1 & 1 \end{bmatrix} x_k + v_k
\end{aligned}
$$

where $w_k \sim (0, Q)$ and $Q = I$.

a) Find one matrix square root of Q.
b) Is (F, H) observable?
c) Is (F, H) detectable?
d) Is (F, G) controllable for all G such that $GG^T = Q$?
e) Is (F, G) stabilizable for all G such that $GG^T = Q$?
f) Use the above results to specify how many positive definite solutions exist to the DARE that is associated with the Kalman filter for this problem.
g) Use the above results to specify whether or not the steady-state Kalman filter for this system is stable.

9. Recall that the steady-state, zero-input, one-step formulation of the *a posteriori* Kalman filter can be written as

$$
\begin{aligned}
\hat{x}_k^+ &= (I - KH)F\hat{x}_{k-1}^+ + Ky_k \\
\hat{y}_k &= H\hat{x}_k^+
\end{aligned}
$$

Prove that if (F, H) is observable and $(I - HK)$ is full rank, then the Kalman filter in the above equation is an observable system. Hint: $H(I - KH) = (I - HK)H$.

Exercise

10. Consider the system

$$x_{k+1} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_k + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_k$$

$$y_k = \begin{bmatrix} 1 & 0 \end{bmatrix} x_k + v_k$$

where w_k and v_k are uncorrelated zero-mean white noise processes with variances q and R, respectively.

a) Use Anderson's approach to reduced-order filtering to estimate the first element of the state vector. Find steady-state values for $\tilde{P}$, $\tilde{\tilde{P}}$, Σ, $\tilde{\Pi}$, $\tilde{\tilde{\Pi}}$, and P. Find the steady-state gain K of the reduced-order filter.

b) Use the full-order filter to estimate the entire state vector. Find steady-state values for P and K.

c) Comment on the comparison between your answer for P in part (a) and part (b).

11.

Suppose that a Kalman filter is running with

$$F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$H = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$R = 1$$

$$P^+(k) = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

An out-of-sequence measurement from time $(k-1)$ is received at the filter.

a) What was the value of $P^-(k)$?

b) Use the delayed-measurement filter to find the quantities $P_w(k, k-1)$, $P_{xw}(k, k-1)$, $P(k-1, k)$, $P_{xy}(k, k-1)$, and $P(k, k-1)$.

c) Realizing that the measurement at time $(k-1)$ was not received at time $(k-1)$, derive the value of $P^-(k-1)$. Now suppose that the measurement was received in the correct sequence at time $(k-1)$. Use the standard Kalman filter equations to compute $P^+(k-1)$, $P^-(k)$, and $P^+(k)$. How does your computed value of $P^+(k)$ compare with the value of $P(k, k-1)$ that you computed in part (b) of this problem?

Exercise

12. Consider the system

$$
\begin{aligned}
x_k &= \frac{1}{2}x_{k-1} + w_{k-1} \\
y_k &= x_k + v_k
\end{aligned}
$$

Note that this is the system model for the radiation system described in Problem 5.1.

a) Find the steady-state value of P_k for the H_∞ filter, using a variable θ and $L = R = Q = S = 1$.

b) Find the bound on θ such that the steady-state H_∞ filter exists.

13. Suppose that you use a continuous-time H_∞ filter to estimate a constant on the basis of noisy measurements. The measurement noise is zero-mean and white with a covariance of R. Find the H_∞ estimator gain as a function of P_0, R, θ, and time. What is the limit of the estimator gain as $t \to \infty$? What is the maximum value of θ such that the H_∞ estimation problem has a solution? How does the value of θ influence the estimator gain?

$$
\begin{aligned}
\Sigma^{-1} &= \tilde{P}^{-1} - \theta L^T L + H^T H \\
\tilde{P} &= F\tilde{P}(H^T H\tilde{P} - \theta L^T L\tilde{P} + I)^{-1}F^T + I \\
\tilde{K} &= (I + \theta L^T L)^{-1}\Sigma H^T \\
&= \tilde{P}(I + H^T H\tilde{P})^{-1}H^T \\
\hat{x}_{k+1} &= F\hat{x}_k + \tilde{K}(y_{k+1} - HF\hat{x}_k)
\end{aligned}
$$

14. Prove that the solution of the *a posteriori* H_∞ Riccati equation given in Equation (11.132) with $\theta = 0$ is equivalent to the solution of the steady-state *a priori* Kalman filter Riccati equation with $R = I$ and $Q = I$.

Exercise

$$\begin{aligned}
V_k &= D_k^T D_k \\
\Sigma_k &= (P_k H_k^T H_k - P_k G_k^T G_k + I)^{-1} P_k \\
P_{k+1} &= (I - V_{k+1}) F_k \Sigma_k F_k^T (I - V_{k+1}) + Q_k \\
K_k &= (I - V_{k+1}) F_k \Sigma_k H_k^T \\
\hat{x}_{k+1} &= F_k \hat{x}_k + K_k (y_k - H_k \hat{x}_k)
\end{aligned}$$

$$(I - G_k P_k G_k^T) \geq 0$$

15. Consider a constrained H_∞ state estimation problem with

$$\begin{aligned}
F &= \begin{bmatrix} 1 & 1/2 \\ 1/2 & 0 \end{bmatrix} \\
G = H &= \begin{bmatrix} G_1 & 0 \end{bmatrix} \\
D &= \begin{bmatrix} 1 & 1 \end{bmatrix} \\
Q &= \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix}
\end{aligned}$$

Find the steady-state constrained Riccati solution for P from Equation (12.50). For what values of G_1 will the condition of Equation (12.51) be satisfied?

16. Suppose you have a scalar system given as

$$\begin{aligned}
x_{k+1} &= x_k \\
y_k &= x_k^2 + v_k
\end{aligned}$$

where v_k is white Gaussian noise with a variance of 0.01. The pdf of the initial state x_0 is uniform between -1 and $+1$. Note from the measurement equation that there is no way to distinguish between a positive state and a negative state.

a) What will the extended Kalman filter estimate of the system be equal to?

b) The pdf of x_0 can be approximated with two Gaussian pdfs, each with a variance of 0.43, and with respective means of $-1/3$ and $+1/3$. Suppose that $x_0 = -1/2$. Plot the true state and the individual state estimates of a two-term Gaussian sum filter for 20 time steps. Plot the Gaussian pdfs at the final time for each estimate of the two-term Gaussian sum filter.

~ THE END ~