# Lecture 25

Last time:

$$K_2 = P^- H^T \left( H P^- H^T + R \right)^{-1}$$

If $\underline{z}_k = \underline{x} + \underline{v}_k$,

$$H = I$$

$$K_2 = P^- \left( P^- + R \right)^{-1}$$

Then if $P^- \ll R$,

$$K_2 \to P^- R^{-1}$$

Alternatively, if $R \ll P^-$,

$$K_2 \to I$$

The effect is that

$$\underline{\hat{x}}^+ = \underline{\hat{x}}^- + K \left( \underline{z} - H \underline{\hat{x}}^- \right)$$

The quantity $K_2$ is known as the <u>Kalman gain</u>. It is the optimum gain in the mean squared error sense. Substitute it into the expression for $P^+$.

$$
\begin{aligned}
P^+ &= \left( I - KH \right) P^- \left( I - KH \right)^T + KRK^T \\
&= \left( I - KH \right) P^- - \left( I - KH \right) P^- H^T K^T + KRK^T \\
&= \left( I - KH \right) P^- - P^- H^T K^T + KHP^- H^T K^T + KRK^T \\
&= \left( I - KH \right) P^- - P^- H^T K^T + K \left( HP^- H^T + R \right) K^T \\
&= \left( I - KH \right) P^- - P^- H^T K^T + P^- H^T \left( HP^- H^T + R \right)^{-1} \left( HP^- H^T + R \right) K^T \\
&= \left( I - KH \right) P^- - P^- H^T K^T + P^- H^T K^T \\
&= \left( I - KH \right) P^-
\end{aligned}
$$

The form at the <u>top</u> is true for <u>any choice of</u> $K$. The <u>last form</u> is true only for the <u>Kalman gain</u>.

The first form is better behaved numerically if you process a <u>measurement</u> which is <u>very accurate</u> relative to your prior information.

So the measurement update step is:

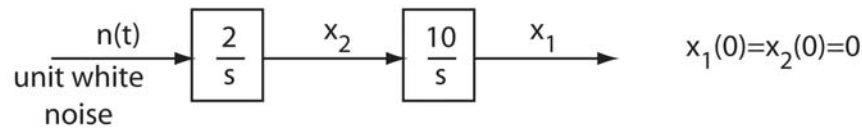$$K = P^-H^T\left(HP^-H^T + R\right)^{-1}$$

$$\underline{\hat{x}}^+ = \underline{\hat{x}}^- + K\left(\underline{z} - H\underline{\hat{x}}^-\right)$$

$$P^+ = \left(I - KH\right)P^-$$

The filter operates by alternating time propagation and measurement update steps.

The results were derived here on the basis of preserving zero mean errors and minimizing the error variances. If all errors and noises are assumed <u>normally distributed</u>, so the probability density functions can be manipulated, one can derive the same results using the <u>conditional mean approach</u>: define $\underline{x}$ at every stage to be the mean of the distribution of $\underline{x}$ conditioned on all the measurements available up to that stage.

---

*Example: Conversion of continuous dynamics to discrete time form*



$$\dot{x}_1 = 10x_2$$

$$\dot{x}_2 = 2n$$

$$\underline{\dot{x}} = \underbrace{\begin{bmatrix} 0 & 10 \\ 0 & 0 \end{bmatrix}}_{A}\underline{x} + \underbrace{\begin{bmatrix} 0 \\ 2 \end{bmatrix}}_{B}n$$

We could do time propagation by integration $(N = 1)$:

$$\dot{\underline{\hat{x}}} = A\underline{\hat{x}}$$

$$\dot{P} = AP + PA^T + BB^T$$

$$\Phi = e^{A\Delta t} = I + A\Delta t + \frac{1}{2}A^2\left(\Delta t\right)^2 + ...$$

$$A^2 = \begin{bmatrix} 0 & 10 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 10 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

If this does not work, expand $\dot{\phi} = A\phi$, $\phi(0) = I$ $(\Delta t = 2)$:

---

$$\Phi_k = I + A\Delta t = \begin{bmatrix} 1 & 10\Delta t \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 20 \\ 0 & 1 \end{bmatrix}$$

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

$$Q_k = \int_0^2 \Phi(2,\tau) BNB^T \Phi(2,\tau)^T \, d\tau$$

$$\Phi(2,\tau) = I + A(2-\tau)$$

$$= \begin{bmatrix} 1 & 10(2-\tau) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 20-10\tau \\ 0 & 1 \end{bmatrix}$$

$$\Phi BNB^T \Phi^T = \begin{bmatrix} 40-20\tau \\ 2 \end{bmatrix} \begin{bmatrix} 40-20\tau & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1600 - 1600\tau + 400\tau^2 & 80 - 40\tau \\ 80 - 40\tau & 4 \end{bmatrix}$$

$$\int_0^2 \Phi BNB^T \Phi^T \, d\tau = \begin{bmatrix} \dfrac{3200}{2} & 80 \\ 80 & 8 \end{bmatrix} = Q_k$$

$$R_k = 4$$

The initial conditions are given to be 0, so

$$\hat{\underline{x}}^-(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$P^-(0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Using the discrete formulation:

Time propagation:

$$\hat{\underline{x}}_{k+1}^- = \Phi \hat{\underline{x}}_k^+$$

$$P_{k+1}^- = \Phi P_k^+ \Phi^T + Q$$

Measurement update:

$$K = P^- H^T \left( H P^- H^T + R \right)^{-1}$$

$$\hat{\underline{x}}^+ = \hat{\underline{x}}^- + K\left( z - H\hat{\underline{x}}^- \right)$$

$$P^+ = \left( I - KH \right) P^-$$

## Continuous time filter

As noted in the beginning of this section, we rarely process measurements continuously in a real system. However, there may be cases in which the measurements are processed so rapidly that one can approximate the process as integration of a continuous time differential equation.
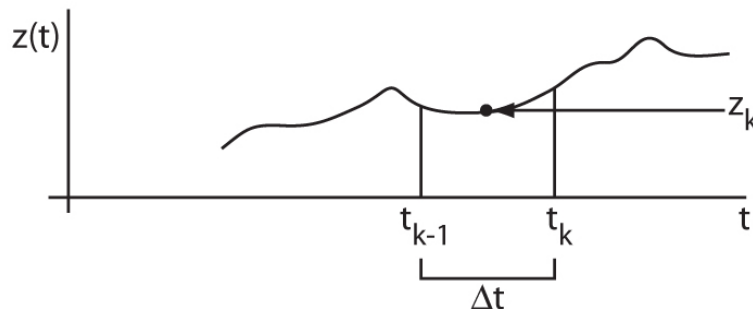
Also, since integration of a single differential equation for the estimate and one for the covariance matrix is logically simpler than alternating time propagation and measurement update steps, we sometimes approximate the hybrid continuous-discrete filter with a nearly equivalent continuous filter for analysis purposes.

The relationship between continuous and discrete measurement processing can be seen in the following. Suppose we had the continuous measurement

$$\underline{z}(t) = H(t)\underline{x}(t) + \underline{v}(t)$$

where $\underline{v}(t)$ is an unbiased white noise process. Define a nearly equivalent discrete measurement to be the average of the continuous measurement over an interval.

$$E\left[\underline{v}(t)\underline{v}(\tau)^{T}\right] = R(t)\delta(t - \tau)$$



$$\underline{z}_k = \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k} \underline{z}(t)d\tau$$

$$= \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k} \left[H(t)\underline{x}(t) + \underline{v}(t)\right]dt$$

$$\approx \frac{1}{\Delta t}H(t_k)\underline{x}(t_k)\Delta t + \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k} \underline{v}(t)dt$$

$$= H_k\underline{x}(t_k) + \frac{1}{\Delta t}\int_{t_{k-1}}^{t_k} \underline{v}(t)dt$$

for small $\Delta t$. The equivalent discrete measurement noise is

$$\underline{v}_k = \frac{1}{\Delta t} \int_{t_{k-1}}^{t_k} \underline{v}(t)dt$$

with $\overline{\underline{v}_k} = 0$ since $\overline{\underline{v}(t)} = 0$, and its covariance is

$$R_k = E\left[\underline{v}_k \underline{v}_k^T\right]$$

$$= \frac{1}{\Delta t^2} \int_{t_{k-1}}^{t_k} dt_1 \int_{t_{k-1}}^{t_k} dt_2 \overline{\underline{v}(t_1)\underline{v}(t_2)^T}$$

$$= \frac{1}{\Delta t^2} \int_{t_{k-1}}^{t_k} dt_1 \int_{t_{k-1}}^{t_k} dt_2 R(t_1)\delta\left(t_2 - t_1\right)$$

$$= \frac{1}{\Delta t^2} \int_{t_{k-1}}^{t_k} dt_1 R(t_1)$$

$$\approx \frac{1}{\Delta t^2} R(t_k)\Delta t$$

$$= \frac{R(t_k)}{\Delta t}$$

for small $\Delta t$.

So when we approximate a discrete measurement by a continuous one, the power density to assign to the continuous measurement noise is
$$R(t) = R_k \Delta t$$

This makes sense because for larger $\Delta t$ (the discrete update process uses fewer measurements) the intensity of the continuous measurement noise is larger (the measurements are not as good).

The discrete measurement gain is
$$K_k = P^- H^T \left(H P^- H^T + R_k\right)^{-1}$$

Using the relation between $R_k$ and $R(t)$ we have

$$K_k = P^- H^T \left(H P^- H^T + \frac{R(t)}{\Delta t}\right)^{-1}$$

Note the difference in units:
Discrete: $\hat{\underline{x}}^+ = \ldots K_k\left(\underline{z}_k - H\hat{\underline{x}}^-\right)$

Continuous: $\dot{\hat{\underline{x}}} = A\hat{\underline{x}} + G\underline{u} + K\left(\underline{z}_k - H\hat{\underline{x}}\right)$

Then the continuous Kalman gain is

$$K(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} K_k$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} P^- H^T \left( HP^- H^T + \frac{R(t)}{\Delta t} \right)^{-1}$$

$$= \lim_{\Delta t \to 0} P^- H^T \left( \Delta t HP^- H^T + R(t) \right)^{-1}$$

$$= P^- H^T R^{-1}$$

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

The distinction between $P^-$ and $P^+$ disappears as we go to the limit of continuous measurement processing.

The continuous-discrete update relations for $\hat{\underline{x}}$ and $P$ can be analyzed in a similar manner, expanding everything to first order in $\Delta t$ and taking the limit as $\Delta t \to 0$. This is done in sufficient detail in the text. The result is the continuous time form of the Kalman filter.

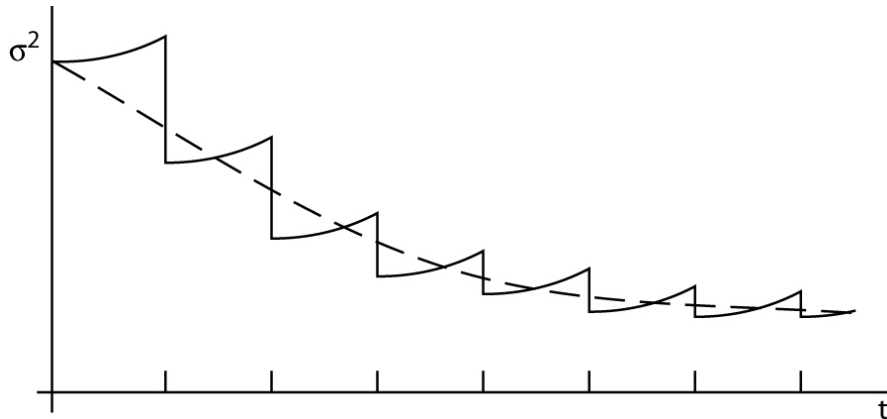$$\dot{\hat{\underline{x}}} = A\hat{\underline{x}} + G\underline{u} + K\left(\underline{z} - H\hat{\underline{x}}\right)$$

$$\dot{P} = AP + PA^T + BNB^T - PH^T R^{-1} HP$$
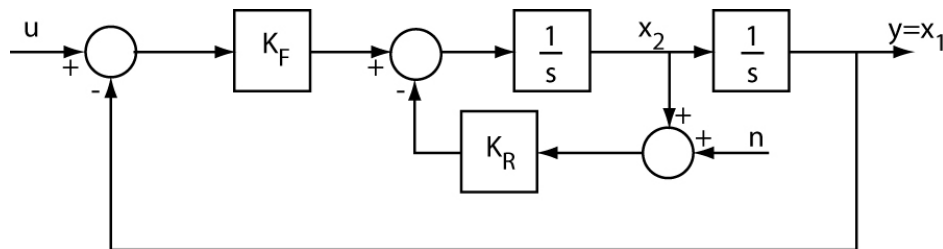
where $K = PH^T R^{-1}$.

If this is an approximation to a filter that processes measurements at discrete points in time, with measurement noise variance $R_k$, take

$$R(t) = R_k \Delta t$$

If the approximation is good, the discrete and continuous variances will be related as

## Example: Discrete time system (typical servo structure)



The input $u(t)$ is any known command signal.

The noise $n(t)$ is a wideband noise approximated as white with power density $N$.

The measurements are of $y$ and are processed at intervals $\Delta t$ with discrete measurement variance $R_k$.

Form the approximating continuous Kalman filter:
$$R = R_k \Delta t$$

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = K_F(u - x_1) - K_R(x_2 + n)$$
$$\underline{\dot{x}} = \begin{bmatrix} 0 & 1 \\ -K_F & -K_R \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ K_F \end{bmatrix} u + \begin{bmatrix} 0 \\ -K_R \end{bmatrix} n$$
$$= A\underline{x} + Gu + Bn$$

The continuous approximation measurement $z$ is

$$z = \begin{bmatrix} 1 & 0 \end{bmatrix} \underline{x} + v$$
$$= H\underline{x} + v$$

The filter is:
$$\dot{\hat{\underline{x}}} = A\hat{\underline{x}} + Gu + K\left(z - H\hat{\underline{x}}\right)$$
$$\dot{P} = AP + PA^T + BNB^T - PH^T R^{-1} HP$$

where $K = PH^T R^{-1}$.

## *Steady state properties*

If the system is time invariant and the random processes are stationary ($A, B, H, N, R$ all constants), then the differential equation for $P$ <u>may</u> approach a constant steady state value as time increases.



Some characteristics of the steady state solution:

1) If $(A, H)$ is <u>detectable</u> and $(A, B)$ is <u>stabilizable</u>, the solution of the Ricatti differential equation approaches a steady state value $P_\infty$ which is <u>independent</u> of the <u>initial condition</u> $P_0$.

2) Under the conditions of (1), this $P_\infty$ is the unique positive semidefinite solution of the algebraic Ricatti equation
$$AP + PA^T + BNB^T - PH^T R^{-1} HP = 0$$

3) The steady state filter
$$\dot{\hat{\underline{x}}} = \left(A - KH\right)\hat{\underline{x}} + G\underline{u} + K\underline{z}$$
is asymptotically stable if and only if $(A, H)$ is detectable at least and $(A, B)$ is stabilizable at least with $K = P_\infty H^T R^{-1}$.

4) Under the conditions of (3), the steady state filter minimizes
$$\lim_{t\to\infty} E\left[\underline{e}(t)^T M\underline{e}(t)\right]$$
for every $M > 0$. This value is
$$\text{tr}\left[P_\infty M\right]$$

5) $P_\infty$ is positive definite if and only if $(A, B)$ is complete controllable.

6) If $(A, H)$ is observable, steady state $P$ is positive definite.