

## Lecture 24

Last time:

$$\frac{d}{dt}X(t) = A(t)X(t) + X(t)A(t)^T + B(t)N(t)B(t)^T$$

If the system reaches a statistical steady state, the covariance matrix would be constant. The system would have to be invariant,

$$\frac{d}{dt}X(t) = AX + XA^T + BNB^T = 0$$

and the steady state covariance matrix can be solved for from this algebraic equation.

Note that:  $X(t)$  is symmetric.

If  $[A, B]$  is controllable and  $X(t_0) > 0$ ,  $X(t) > 0$  for all  $t$ .

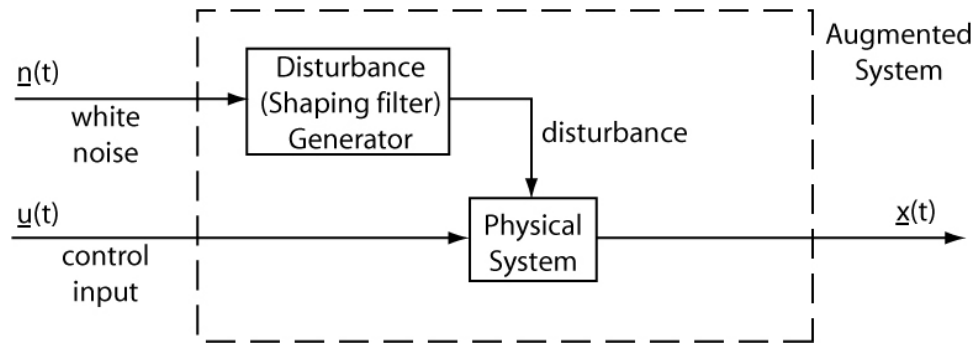
If, in addition,  $[A, C]$  is observable, then  $\text{Cov}[\underline{y}(t)] > 0$  for all  $t$ .

## Kalman Filter

Our text treats two forms of the Kalman filter:

- Discrete time filter  
Based on the model
  - System:  $\underline{x}(k+1) = \phi_k \underline{x}(k) + \underline{w}_k$
  - Measurements:  $\underline{z}_k = H_k \underline{x}(k) + \underline{v}_k$
- Continuous time filter
  - System:  $\dot{\underline{x}} = F\underline{x} + G\underline{w}$
  - Measurements:  $\underline{z} = H\underline{x} + \underline{v}$

The most common practical situation is neither of these. In the aerospace area, and many other application areas as well, we are interested in dynamic systems that evolve continuously in time, so we must at least start with a continuous system model. The system may be driven by a known command input and may be subject to random disturbances. If the given disturbances cannot be well approximated as white, then a shaping filter must be added to shape the given disturbance spectrum from white noise.



The augmented system is modeled as

$$\dot{\underline{x}} = A(t)\underline{x} + G(t)\underline{u} + B(t)\underline{n}$$

where  $\underline{n}(t)$  is an unbiased white noise with correlation function

$$E[\underline{n}(t)\underline{n}(\tau)^T] = N(t)\delta(t - \tau)$$

We allow the system matrices to be time varying. The most common reason for dealing with time varying linear dynamics is because we have linearized nonlinear dynamics about some trajectory.

But although the system dynamics are continuous in time, we do not process measurements continuously. The Kalman filter involves significant computations, and those computations are universally executed in a digital computer. Since a digital computer works on a cyclic basis, the measurements are processed at discrete points in time.

The measurements are modeled as

$$\underline{z}_k = H_k \underline{x}(t_k) + \underline{v}_k$$

where  $\underline{v}_k$  is an unbiased, independent noise sequence with

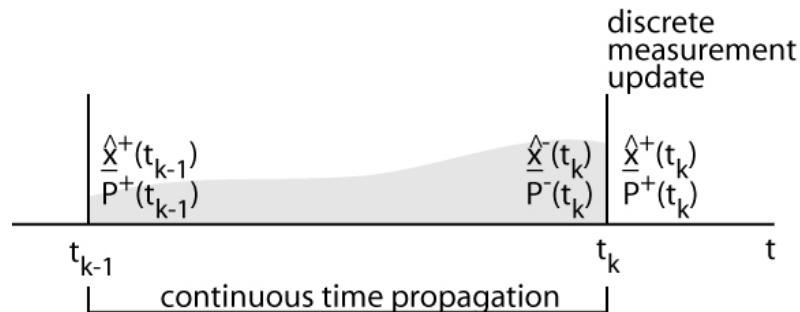
$$E[\underline{v}_k \underline{v}_k^T] = R_k$$

If the measurements have a noise component which is correlated in time, a continuous time shaping filter must be added to the augmented system model designed to give the proper correlation for the noise samples. We assume here that this augmentation, if needed, has been done.

Although we can allow for a correlated measurement noise component, it is essential that there be an independent noise component as well. If there is no independent noise component, the whole nature of the optimum filter changes.

So the physical situation of most common interest is a hybrid one:

- System dynamics are continuous in time
- Measurements are processed at discrete points in time



Where a discrete time measurement is processed, there is a discontinuity in the estimate, the estimate error, and the error covariance matrix. We use the superscript “-” to indicate values before incorporating the measurement at a measurement time and the superscript “+” to indicate values after incorporating the measurement at a measurement time.

The filter operates in a sequence of two-part steps

1. Propagate  $\hat{x}(t)$  and  $P(t)$  in time starting from  $\hat{x}^+(t_{k-1})$  and  $P^+(t_{k-1})$  to  $\hat{x}^-(t_k)$  and  $P^-(t_k)$
2. Incorporate the measurement at  $t_k$  to produce  $\hat{x}^+(t_k)$  and  $P^+(t_k)$

### ***Time Propagation***

After processing the measurement at  $t_{k-1}$  we have the estimate  $\hat{x}^+(t_{k-1})$  and the estimation error covariance matrix  $P^+(t_{k-1})$ .

Then the filter operates open loop until the next measurement point at  $t_k$ . Note that we are not considering any effects of computation delays here.

In the interval from  $t_{k-1}$  to  $t_k$  the system dynamics are described as  

$$\dot{\underline{x}} = A(t)\underline{x} + G(t)\underline{u} + B(t)\underline{n}$$

The estimate  $\hat{x}^+(t_{k-1})$  is the mean of the distribution of  $\underline{x}$  conditioned on all the measurements processed up to that point. So the estimate error is zero mean. We want to preserve the error unbiased until the next measurement, and we can achieve this driving the estimate with the mean or expectation of the system dynamics.

$$\dot{\underline{\hat{x}}} = A(t)\underline{\hat{x}} + G(t)\underline{u}$$

$$\underline{\hat{x}}(t_{k-1}) = \underline{\hat{x}}^+(t_{k-1})$$

The error dynamics are then

$$\underline{e}(t) = \underline{x}(t) - \underline{\hat{x}}(t)$$

$$\dot{\underline{e}} = \underline{\dot{x}} - \dot{\underline{\hat{x}}}$$

$$= A(t)\underline{e} + B(t)\underline{n}$$

The mean of the error then satisfies

$$\dot{\underline{\bar{e}}} = A(t)\underline{\bar{e}}$$

since  $\underline{\bar{n}}(t) = 0$ , and since the initial condition is

$$\underline{\bar{e}}(t_{k-1}) = 0$$

we have  $\underline{\bar{e}}(t) = 0$ ,  $(t_{k-1}^+, t_k^-)$

The form of the error dynamics are a linear system driven by white noise, which is the form of dynamics we just analyzed. The covariance matrix for the estimation error therefore satisfies the differential equation

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

with the initial condition

$$P(t_{k-1}) = P^+(t_{k-1})$$

So the time propagation step can be taken by integrating the differential equations

$$\dot{\underline{\hat{x}}} = A\underline{\hat{x}} + G\underline{u} \quad \underline{\hat{x}}(t_{k-1}) = \underline{\hat{x}}^+(t_{k-1})$$

$$\dot{P} = AP + PA^T + BNB^T \quad P(t_{k-1}) = P^+(t_{k-1})$$

Both of these relations can be expressed in the form of discrete time transitions, but especially if  $A$  and/or  $B$  and/or  $G$  are time-varying, this may not be convenient.

The transition matrix would be the solution to the differential equation

$$\frac{d}{dt}\Phi(t, t_{k-1}) = A(t)\Phi(t, t_{k-1})$$

with the initial condition

$$\Phi(t_{k-1}, t_{k-1}) = I$$

Then the estimate at the next measurement point would be given by

$$\hat{\underline{x}}^-(t_k) = \Phi(t_k, t_{k-1}) \hat{\underline{x}}^+(t_{k-1}) + \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) G(\tau) \underline{u}(\tau) d\tau$$

Note the simplification if  $\underline{u}(\tau)$  is constant over this interval. The corresponding covariance matrix would be

$$P^-(t_k) = \Phi(t_k, t_{k-1}) P^+(t_{k-1}) \Phi(t_k, t_{k-1})^T + \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) B(\tau) N(\tau) B(\tau)^T \Phi(t_k, \tau)^T d\tau$$

Both of these expressions are much more complex to deal with than would be the case if one were just given a discrete time model of the system dynamics - such as is done in the text.

$$\underline{x}_{k+1} = \phi_k \underline{x}_k + G_k \underline{u}_k + \underline{w}_k$$

$$\hat{\underline{x}}_{k+1}^- = \phi_k \hat{\underline{x}}_k^+ + G_k \underline{u}_k$$

$$P_{k+1}^- = \phi_k P_k^+ \phi_k^T + Q_k$$

where  $Q_k = \overline{\underline{w}_k \underline{w}_k^T}$

If you wanted to express continuous time dynamics in discrete time form, you could only do it if  $\underline{u}(\tau)$  were constant ( $= \underline{u}_k$ ) in  $(t_k, t_{k+1})$ . Then

$$\phi_k = \Phi(t_{k+1}, t_k)$$

$$G_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) G(\tau) d\tau$$

$$\underline{w}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) B(\tau) \underline{n}(\tau) d\tau$$

$$Q_k = \overline{\underline{w}_k \underline{w}_k^T}$$

$$= \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) B(\tau) N(\tau) B(\tau)^T \Phi(t_{k+1}, \tau)^T d\tau$$

### Measurement Update

At the next measurement point,  $t_k$ , we have the a priori estimate,  $\hat{\underline{x}}^-(t_k)$ , its error covariance matrix,  $P^-(t_k)$ , and a new measurement of the form

$$\underline{z}_k = H_k \underline{x}(t_k) + \underline{v}_k$$

where the noises,  $\underline{v}_k$ , are an unbiased, independent sequence of random variables with covariance matrix,  $R_k$ .

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

We wish to incorporate the information contained in this measurement in our estimate of the state.

To this end, postulate an updated estimate,  $\hat{\underline{x}}^+$ , in the form of a general linear combination of the a priori estimate and the measurement. For simplicity I will suppress the time arguments.

$$\hat{\underline{x}}^+ = K_1 \hat{\underline{x}}^- + K_2 \underline{z}$$

We wish to choose  $K_1$  and  $K_2$  so as to:

1. Preserve the estimation error unbiased
2. Minimize the a posteriori error variance

The a posteriori error is

$$\begin{aligned} \underline{e}^+ &= \underline{x} - \hat{\underline{x}}^+ \\ &= \underline{x} - K_1 \hat{\underline{x}}^- - K_2 \underline{z} \\ \underline{e}^+ &= \underline{x} - K_1 (\underline{x} - \underline{e}^-) - K_2 (H\underline{x} + \underline{v}) \\ &= (I - K_1 - K_2 H) \underline{x} + K_1 \underline{e}^- - K_2 \underline{v} \end{aligned}$$

The mean a posteriori error is (remember that the a priori error and the measurement noise are zero-mean)

$$\overline{\underline{e}^+} = (I - K_1 - K_2 H) \underline{x}$$

We want this to be zero for all  $\underline{x}$ . Therefore

$$K_1 = I - K_2 H$$

With this choice of  $K_1$ ,

$$\overline{\underline{e}^+} = 0$$

and

$$\begin{aligned} \hat{\underline{x}}^+ &= (I - K_2 H) \hat{\underline{x}}^- + K_2 \underline{z} \\ &= \hat{\underline{x}}^- + K_2 (\underline{z} - H \hat{\underline{x}}^-) \end{aligned}$$

So the updated estimate is the a priori estimate plus a gain times the measurement residual. We observed this form earlier when discussing estimation in general.

Now we wish to find the  $K_2$  that minimizes the a posteriori error variance.

$$\begin{aligned}\hat{\underline{x}}^+ &= \hat{\underline{x}}^- + K_2 (H\underline{x} + \underline{v} - H\hat{\underline{x}}^-) \\ &= \hat{\underline{x}}^- + K_2 (H\underline{e}^- + \underline{v}) \\ \underline{e}^+ &= \underline{x} - \hat{\underline{x}}^+ \\ &= \underline{x} - \hat{\underline{x}}^- - K_2 (H\underline{e}^- + \underline{v}) \\ &= (I - K_2 H) \underline{e}^- - K_2 \underline{v}\end{aligned}$$

Note that  $\underline{v}$  is independent of  $\underline{e}^-$ .

$$\begin{aligned}P^+ &= E[\underline{e}^+ \underline{e}^{+T}] \\ &= (I - K_2 H) \overline{\underline{e}^- \underline{e}^{-T}} (I - K_2 H)^T + K_2 \overline{\underline{v} \underline{v}^T} K_2^T \\ &= (I - K_2 H) P^- (I - K_2 H)^T + K_2 R K_2^T\end{aligned}$$

Our optimization criterion will be taken to be the sum of the variances of the a posteriori estimation errors. The sum of the error variances is the trace of  $P^+$

$$\begin{aligned}\text{Min } J &= \text{tr}(P^+) \\ &= \text{tr}\left[(I - K_2 H) P^- (I - K_2 H)^T\right] + \text{tr}\left[K_2 R K_2^T\right]\end{aligned}$$

We use a derivative form

$$\frac{\partial}{\partial A} \left[ \text{tr}(ABA^T) \right] = 2AB$$

if  $B$  is symmetric.

$$\begin{aligned}\frac{\partial J}{\partial K_2} &= 2(I - K_2 H) P^- \frac{\partial}{\partial K_2} (I - K_2 H)^T + 2K_2 R \\ &= 2\left\{ (I - K_2 H) P^- (-H^T) + K_2 R \right\} \\ &= 2\left\{ -P^- H^T + K_2 (HP^- H^T + R) \right\} = 0 \\ K_2 &= P^- H^T (HP^- H^T + R)^{-1}\end{aligned}$$

This is known as the Kalman gain.