Lecture Notes #13.3 Continuous Time Filtering

1. Model:

\[ x = Ax + Bw \quad ; \quad w(t) \sim \text{GWN}(0, Q \delta(t-\tau)) \quad , \quad x(t_0) \sim \mathcal{N}(\hat{x}_0, P_0) \]

\[ z = Hx + v \quad ; \quad v(t) \sim \text{GWN}(0, R \delta(t-\tau)) \quad w, v, x_0 \text{ mutually independent} \] (1)

2. Filter (Estimator):

\[ \hat{x} = A\hat{x} + P^{\top}R^{-1}(z - H\hat{x}) \quad ; \quad \hat{x}(t_0) = \hat{x}_0 \]

\[ P = AP + PA^T + BQ^T R^{-1} HP \quad ; \quad P(t_0) = P_0 \] (2)

where \( z - H\hat{x} \equiv \eta \) is the innovation process which behaves as a \( \text{GWN}(0, R \delta(t-\tau)) \). This is as the case should be since \( \eta(t) \) being a white noise represents the totally new and independent information that is conveyed by \( z \).

3. Now consider the special case of \( Q = 0 \) (no disturbance) \( R = I \) (for notational simplicity) and \( P(t_0) = \) (no prior knowledge whatsoever), we have

\[ P = AP + PA^T + BQ^T HP \quad ; \quad P(t_0) = \cdot \]

Using the trick \( P^{-1}P = I \Rightarrow P^{-1}P + P^{-1}P = 0 \), we can derive

\[ \frac{dP^{-1}}{dt} = -P^{-1}A^TP^{-1} + H^TH \quad ; \quad P^{-1}(t_0) = 0 \]

with the solution

\[ P^{-1}(t) = \int_{t_0}^{t} \Phi^T(t_0, \tau)H^TH\Phi(t_0, \tau)d\tau \] (3)

The condition that \( P^{-1}(t) > 0 \) for some \( t \), then insure the existence of \( P(t) \) which is equivalent to the condition of observability of the system \( dx/dt = Ax \) and \( z = Hx \). This makes total intuitive sense, namely, only if the deterministic system is observable, will we have a chance of getting finite variance from infinite variance via filtering.

If we further specialize to the case of scalar \( x \), \( A = 0 \), \( H = 1 \), then we have the model of \( dx/dt = 0 \), and \( z = x + v \), the \( P \) equation becomes

\[ P = -P^2 \text{ or } P^{-1} = 1 \quad ; \quad P(t) = \frac{1}{t} \quad \text{ and } \quad \hat{x} = \frac{1}{t}(z - \hat{x}) \]

We claim a solution of the \( \hat{x} \) equation is

\[ \hat{x}(t) = \frac{1}{t} \int_{t_0}^{t} z(\tau)d\tau = x + \frac{1}{t} \int_{t_0}^{t} v(\tau)d\tau \]

(Check: \( \hat{x} = -\frac{1}{t} \int_{0}^{t} z(t)d\tau + \frac{1}{t^2} \int_{0}^{t} z(t)dt \) equals \( \frac{1}{t}(z - \hat{x}) \) !)

Thus in the simplest case, Kalman filter is nothing but the recursive implementation of averaging out the noise.

4.
\[
\Phi^{T-1-i} H^{T} \Phi^{T-1-i} > 0
\]

for the system \( x_{t+1} = \Phi x_t, z_t = H x_t \).

**Miscellaneous Issue Associated with Filtering**

5. It should also be clear that if the model of §1 is modified to become,

\[
x = Ax + Bw + Gu; \quad w(t) \sim \text{GWN}(0, Q\delta(t-\tau)), \quad x(t_0) \sim \text{N}(\hat{x}_0, P_0) \tag{1}'
\]

\[
z = Hx + v; \quad v(t) \sim \text{GWN}(0, R\delta(t-\tau)) \quad \text{w, v, } x_0 \text{ mutually independent}
\]

where the term \( Gu \) is added to represent the control variable, then the only modification to Eq.(2) should be

\[
\hat{x} = A\hat{x} + PH^{T}R^{-1}(z-H\hat{x}) + Gu; \quad \hat{x}(t_0) = \hat{x}_0 \tag{2}'
\]

\[
P = AP + PA^{T} + BQB^{T}PH^{T}R^{-1}HP; \quad P(t_0) = P_0
\]

which is based on the principle that:

*Estimation means that we should model everything that is known. Use the knowledge to subtract out the predictable components of the measurement by building a model of the system. What are left should be used to update the state according to the weighting factor \( PH^{T}R^{-1} \). The weighting represents a trade off between averaging out the noise and accounting for estimation errors.*

6. If we define \( e(t) \equiv x(t) - \hat{x}(t) \). It is directly verified that

\[
e = (A-PH^{T}R^{-1}H)e + Bw - PHRv
\]

which is independent of the control. This means there are irreducible errors in stochastic control that cannot be effected by clever control laws.

7. Consider the generalized least square fit problem of determine optimal \( x(\tau) \) and \( w(\tau) \) for \( t_0 \leq \tau \leq t \), denoted as \( \hat{x}(\tau) \), to minimize

\[
J = \frac{1}{2} \| x(t_0) - \hat{x}_0 \|_{P_0}^2 + \frac{1}{2} \int_{t_0}^{t} \left[ \| z_\tau - Hx \|_{R_\tau}^2 + \| w \|_{Q_\tau}^2 \right] dt
\]

subject to \( \dot{x} = Ax + Bw \)

At time \( t \), \( \hat{x}(\tau) \) is the best estimate corresponding to the filtering estimate given by (2) under probabilistic assumptions. For \( \tau < t \), \( \hat{x}(\tau) \) is the "smoothed" estimate of \( x(t) \) given the history \( z(\tau), t_0 \leq \tau \leq t \).

8. If \( w \) and \( v \) are not GWN, we can always model them as Gauss-Markov processes, i.e., linear dynamic systems driven by white Gaussian noise. We'll be back to the model of (1) again. The possible complication is that we may have the situation where

\[
z = Hx_{\text{enlarged}}
\]

We need some small modification to handle this which is outside the scope of this course.
9. Finally, the accurate numerical solution of the P equation in continuous and discrete time is a subject in its own right which we shall not treat here.