5 The Continuous-Time Kalman Filter

The Model: Continuous-time linear system, with white noises state and measurement noises (not necessarily Gaussian).

Goal: Develop the continuous-time Kalman filter as the optimal *linear* estimator (L-MMSE) for this system.

One way to develop the continuous-time filter is as the limit (with $\Delta T \to 0$) of the discrete time case. The derivation below follows a direct approach, based on the innovations process, introduced by Kailath, and will be somewhat informal.

A rigorous but very accessible treatment may be found in: M. Davis, *Linear Estimation and Stochastic Control,* 1977. However, this more advanced treatment is only essential in the nonlinear estimation problem.

5.1 The continuous time model

We consider the state-space model

$$\frac{d}{dt}x_t = F_t x_t + w_t\,, \quad t \geq 0$$
$$z_t = H_t x_t + v_t$$

where:

- $\{w_t\}$ and $\{v_t\}$ are zero-mean white-noise processes, namely

$$E(w_t w_s^T) = Q_t \delta(t-s), \quad E(v_t v_s^T) = R_t \delta(t-s)$$
$$E(w_t v_s^T) = 0$$

- $\{w_t\}$, $\{v_t\}$ and x_0 are uncorrelated.
- $E(x_0) = \overline{x}_0$ and $\mathrm{cov}(x_0) = P_0$ are given.
- We shall assume that R_t is non-singular.

A simplifying assumption: We assume in the derivation below that $\overline{x}_0 = 0$, hence $\overline{x}_t \doteq E(x_t) = 0$. Otherwise, $\overline{x}_t$ is given by $\dot{\overline{x}}_t = F_t \overline{x}_t$ and needs to be added in some of the intermediate equations. The filter equations are the same.

Remark: The white processes above are not rigorously defined, due to the δ-covariances, and indeed their sample-paths are quite "hectic". A rigorous definition of such processes (and the above model) is based on their *integral* — e.g., a Brownian motion in the Gaussian case.

State Covariance Propagation: $\Pi_t := \mathrm{cov}(x_t)$ satisfies

$$\frac{d}{dt}\Pi_t = F_t\Pi_t + \Pi_t F_t^T + Q_t \tag{1}$$

with $\Pi_0 = \mathrm{cov}(x_0)$.

Proof: A naive proof approach by differentiating $E(x_t x_t^T)$ inside the expectation runs into trouble, because of the unusual properties of w_t. The following two options lead to the correct answer:

1. As a limit of the discrete-time case, with the approximation:
   $x_{(k+1)\epsilon} = (I + \epsilon F_{k\epsilon})x_{k\epsilon} + w_{k\epsilon}$, with $E(w_{k\epsilon}w_{l\epsilon}) = \epsilon Q_{k\epsilon}\delta_{kl}$, and $\epsilon \to 0$.
2. Explicitly solving for x_t: Let $\Phi(t,s)$ denote the $n \times n$ state transition matrix, namely the unique solution (for each s) of

   $$\frac{d}{dt}\Phi(t,s) = F_t\Phi(t,s), \quad \Phi(s,s) = I \tag{2}$$

   Then

   $$x_t = \Phi(t,0)x_0 + \int_0^t \Phi(t,s)w_s ds\,. \tag{3}$$

   This can be used to derive the covariance equation. □

Example: Consider the stationary case – the system and covariance matrices are time independent.
Then the white noise processes have constant spectral densities:

$$S_w(\omega) = Q, \quad S_v(\omega) = R$$

and the noise-to state transfer function is

$$T(s) := T_{w\to x}(s) = (sI - F)^{-1}\,.$$

When this system is stable, the state spectral density is given by $S_x(\omega) = T(j\omega)QT^*(j\omega)$, and the measurement spectral density is $S_z(\omega) = HS_x(\omega)H^T + R$.

5.2 Filter Derivation

Let $Z_t = \{z_s,\ s < t\}$. We need to calculate $\hat{x}_t = E^L(x_t|Z_t)$.

Define the innovations process:

$$\tilde{z}_t := z_t - E^L(z_t|Z_t) \tag{4}$$

Observe that

$$\tilde{z}_t = z_t - H_t E^L(x_t|Z_t) = H_t\tilde{x}_t + v_t \tag{5}$$

where $\tilde{x}_t = x_t - \hat{x}_t$.

<u>Properties of $\tilde{z}_t$</u>:

$\tilde{z}_t$ is a zero-mean *white* noise process (exercise). Its covariance equals:

$$E(\tilde{z}_t\tilde{z}_s^T) = R_t\delta(t-s)\,. \tag{6}$$

Note: same covariance as z_t !

It can also be shown that $\tilde{Z}_t$ and Z_t are linearly equivalent, so that

$$E^L(\cdot|Z_t) = E^L(\cdot|\tilde{Z}_t)\,.$$

It follows that $\hat{x}_t$ can be expressed as a linear function of $\tilde{Z}_t$:

$$\hat{x}_t = \int_0^t g(t,s)\tilde{z}_s ds \tag{7}$$

The kernel $g(t,s)$ is easily computable via the orthogonality principle. Since $\tilde{x}_t := (x_t - \hat{x}_t) \perp \tilde{z}_s$ for $s < t$,

$$\begin{aligned} E(x_t\tilde{z}_s^T) &= \int_0^t g(t,r)E(\tilde{z}_r\tilde{z}_s^T)dr \\ &= \int_0^t g(t,r)\delta(s-r)R_s dr = g(t,s)R_s, \quad s < t\,. \end{aligned}$$

Therefore,

$$\hat{x}_t = \int_0^t E(x_t \tilde{z}_s^T) R_s^{-1} \tilde{z}_s ds \tag{8}$$

Differentiate to obtain a differential equation for $\hat{x}$:

$$\frac{d}{dt}\hat{x}_t = \int_0^t E(\dot{x}_t \tilde{z}_s^T) R_s^{-1} \tilde{z}_s ds + E(x_t \tilde{z}_t^T) R_t^{-1} \tilde{z}_t \tag{9}$$

Now, define $K_t = E(x_t \tilde{z}_t^T) R_t^{-1}$, substitute $\dot{x}_t$ from the state equation, note (8), and that $E(w_t \tilde{z}_s^T) = 0$ for $s < t$. These give

$$\frac{d}{dt}\hat{x}_t = F_t \hat{x}_t + K_t \tilde{z}_t = F_t \hat{x}_t + K_t (z_t - H_t \hat{x}_t)\,. \tag{10}$$

Using (5)

$$K_t := E(x_t \tilde{z}_t^T) R_t^{-1} = E(x_t \tilde{x}_t^T) H_t^T R_t^{-1} + 0 = P_t H_t^T R_t^{-1} \tag{11}$$

where $P_t := E(\tilde{x}_t \tilde{x}_t^T)$.

<u>Calculating P_t:</u>

Recall that $\tilde{x}_t = x_t - \hat{x}_t$. Using (10),

$$\frac{d}{dt}\tilde{x}_t = (F_t - K_t H_t)\tilde{x}_t + w_t - K_t v_t\,. \tag{12}$$

As in (1), this implies

$$\frac{d}{dt}P_t = (F_t - K_t H_t) P_t + P_t (F_t - K_t H_t)^T + Q_t + K_t R_t K_t^T \tag{13}$$

with P_0 given.

An alternative expression: by substituting K_t from (11) and rearranging,

$$\frac{d}{dt}P_t = F_t P_t + P_t F_t^T + Q_t - K_t R_t K_t^T\,. \tag{14}$$

This is the (differential) *Riccati Equation* - a quadratic matrix differential equation.

To summarize: The filter equation is given by (10)

$$\frac{d}{dt}\hat{x}_t = F_t\hat{x}_t + K_t\tilde{z}_t = F_t\hat{x}_t + K_t(z_t - H_t\hat{x}_t)$$

with the gain (11)

$$K_t := E(x_t\tilde{z}_t^T)R_t^{-1} = P_tH_t^TR_t^{-1}\,.$$

The covariance may be computed by (the Joseph form)

$$\frac{d}{dt}P_t = (F_t - K_tH_t)P_t + P_t(F_t - K_tH_t)^T + Q_t + K_tR_tK_t^T$$

or (14)

$$\frac{d}{dt}P_t = F_tP_t + P_tF_t^T + Q_t - K_tR_tK_t^T\,.$$

Note: if the state and measurement noises are correlated, namely

$$E(w_tv_s^T) = S_t\delta(t-s)\,,$$

then the gain in (11) becomes $K_t = (P_tH_t^T + S_t)R_t^{-1}$, and the covariance update (13) should be modified by adding $-(S_tK_t^T + K_tS_t^T)$ on the right.