Derivation of Kalman filter equations

$X_t \in \mathbb{R}^m$
$Y_t \in \mathbb{R}^k$

$$\begin{cases} X_{t+1} = F_t X_t + G_t U_t & \text{state equation} \quad (1) \\ Y_t = H_t X_t + V_t & \text{measurement equation (output)} \quad (2) \end{cases}$$

where $\begin{cases} t = 0, 1, 2, \cdots; \; X_t = 0 \text{ for } t < 0; \quad E(X_0) = \bar{X}_0, \; Cov(X_0) = \Sigma_0, \\ \qquad E(U_t) = 0, \; Cov(U_t) = Q_t, \\ \qquad E(V_t) = 0, \; Cov(V_t) = R_t \\ X_0 \perp U_0 \perp U_1 \perp \cdots \perp V_0 \perp V_1 \perp \cdots \\ F_t, G_t, H_t \text{ and the statistics } \bar{X}_0, \Sigma_0, Q_t, R_t \text{ are known} \end{cases}$

<u>Problem</u> : Find $\hat{X}_{t|t} \triangleq \hat{E}(X_t \mid Y_0^t)$

$$\hat{X}_{t+1|t} \triangleq \hat{E}(X_{t+1} \mid Y_0^t)$$

<u>Solution</u>

- Apply $\hat{E}(\cdot \mid Y_0^t)$ to state equation (1)

$$\Rightarrow \hat{X}_{t+1|t} = F_t \hat{X}_{t|t} + G_t \cancel{\hat{U}_{t|t}}^{\,0}$$ because $E(U_t) = 0$ and $Y_0^t \in \text{Span}\{X_0, U_0^{t-1}, V_0^t\} \perp U_t$

$$\Rightarrow \boxed{\hat{X}_{t+1|t} = F_t \hat{X}_{t|t}} \qquad \underline{\text{time update}} \qquad (3)$$

- Now evaluate $\hat{X}_{t|t}$ in (3):

$$\hat{X}_{t|t} \triangleq \hat{E}(X_t \mid Y_0^t)$$
$$= \hat{E}(X_t \mid Y_0^{t-1}, Y_t)$$

$$\hat{X}_{t|t} = \hat{E}(X_t \mid Y_0^{t-1}, I_t) \qquad \text{where } \boxed{I_t \triangleq Y_t - \hat{Y}_{t|t-1} \quad \text{innovation}}$$

$$= \underbrace{\hat{E}(X_t \mid Y_0^{t-1})} + \hat{E}(X_t \mid I_t) \qquad \text{because } I_t \perp Y_0^{t-1}$$

$$= \hat{X}_{t|t-1} + \hat{E}(X_t - \hat{X}_{t|t-1} \mid I_t) \qquad \text{because } \hat{X}_{t|t-1} \text{ is a linear fcn of } Y_0^{t-1} \Rightarrow \text{is} \perp I_t$$

$$= \hat{X}_{t|t-1} + \underbrace{Cov(X_t - \hat{X}_{t|t-1}, I_t)}\ \underbrace{Cov^{-1}(I_t)}\ I_t \qquad (4)$$

Need to evaluate both Covariance matrices in (4).

From measurement equation (2), applying $\hat{E}(\bullet \mid Y_0^{t-1})$, get

$$\hat{Y}_{t|t-1} = H_t \hat{X}_{t|t-1} + \cancel{\hat{V}_{t|t-1}}^{\,0} \qquad \text{because } E(V_t) = 0 \text{ and } V_t \perp Y_0^{t-1}$$

$$\Rightarrow \quad I_t = Y_t - \hat{Y}_{t|t-1}$$

$$= H_t (X_t - \hat{X}_{t|t-1}) + V_t \qquad (5)$$

Define $\boxed{\Sigma_{t|t-1} \triangleq Cov(X_t - \hat{X}_{t|t-1})}$ Covariance of prediction error at time t

$\Rightarrow$ Covariance matrices in (4) are given by

$$Cov(X_t - \hat{X}_{t|t-1}, I_t) = Cov(X_t - \hat{X}_{t|t-1}, H_t(X_t - \hat{X}_{t|t-1}) + V_t)$$

$$= Cov(X_t - \hat{X}_{t|t-1}, H_t(X_t - \hat{X}_{t|t-1}))$$

because $V_t \perp U_0^t, V_0^{t-1}$

$$= \Sigma_{t|t-1} H_t^T$$

$$Cov(I_t) = Cov(H_t(X_t - \hat{X}_{t|t-1}) + V_t)$$
$$= H_t \Sigma_{t|t-1} H_t^T + R_t$$

Replacing in (4), we obtain

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + \underbrace{\Sigma_{t|t-1} H_t^T (H_t \Sigma_{t|t-1} H_t^T + R_t)^{-1}}_{} I_t$$

(from (5))

$$\boxed{\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t (Y_t - H_t \hat{X}_{t|t-1})} \quad (6)$$

= Measurement update

where

$$\boxed{K_t = \Sigma_{t|t-1} H_t^T (H_t \Sigma_{t|t-1} H_t^T + R_t)^{-1}} = \text{Kalman Gain} \quad (7)$$

- The time update eqn (3) and the measurement update eqn (6) give an update for $\hat{X}_{t|t}$ and $\hat{X}_{t|t-1}$, supposing that K_t is available:

$$\boxed{\begin{aligned} \hat{X}_{t+1|t} &= F_t \hat{X}_{t|t} & (3) \\ \hat{X}_{t|t} &= \hat{X}_{t|t-1} + K_t (Y_t - H_t \hat{X}_{t|t-1}) & (6) \end{aligned}} \quad t \geq 0$$

With initial conditions $\hat{X}_{0|-1} = E(X_0)$.

- Need to compute & update $\Sigma_{t|t-1}$ in order to get Kalman gain K_t!

We have

$$\Sigma_{t+1|t} \triangleq Cov(X_{t+1} - \hat{X}_{t+1|t})$$
$$= Cov(F_t(X_t - \hat{X}_{t|t}) + G_t U_t)$$

where $\begin{cases} X_{t+1} = F_t X_t + G_t U_t \\ \hat{X}_{t+1|t} = F_t \hat{X}_{t|t} + G_t \cancel{\hat{U}_{t|t}} \end{cases}$ 0

$$\Rightarrow \boxed{\Sigma_{t+1|t} = F_t \Sigma_{t|t} F_t^T + G_t Q_t G_t^T} \quad (8)$$

- Need to derive an expression for $\Sigma_{t|t}$ in (8) now!

$$\Sigma_{t|t} \triangleq Cov(X_t - \hat{X}_{t|t})$$

$$= Cov(X_t - \hat{X}_{t|t} \mid Y_0^t) \qquad \text{because } X_t - \hat{X}_{t|t} \perp Y_0^t$$

$$= Cov(X_t \mid Y_0^t) \qquad \text{because } \hat{X}_{t|t} \text{ is a linear fcn of } Y_0^t$$

$$= Cov(X_t \mid Y_0^{t-1}, I_t)$$

$$= Cov(X_t - \hat{X}_{t|t-1} \mid Y_0^{t-1}, I_t) \qquad \text{because } \hat{X}_{t|t-1} \text{ is a linear fcn of } Y_0^{t-1}$$

$$= Cov(X_t - \hat{X}_{t|t-1} \mid I_t) \qquad \text{because } X_t - \hat{X}_{t|t-1} \perp Y_0^{t-1}$$

$$= \underbrace{Cov(X_t - \hat{X}_{t|t-1})}_{\Sigma_{t|t-1}} - \underbrace{Cov(X_t - \hat{X}_{t|t-1}, I_t)}_{\text{see (4)}} \underbrace{Cov^{-1}(I_t)}_{\text{see (4)}} Cov(I_t, X_t - \hat{X}_{t|t-1})$$

$$= \Sigma_{t|t-1} - \underbrace{\Sigma_{t|t-1} H_t^T (H_t \Sigma_{t|t-1} H_t^T + R_t)^{-1}} H_t \Sigma_{t|t-1}$$

$$\boxed{\Sigma_{t|t} = \Sigma_{t|t-1} - K_t H_t \Sigma_{t|t-1}} \qquad \text{using (7)} \qquad (9)$$

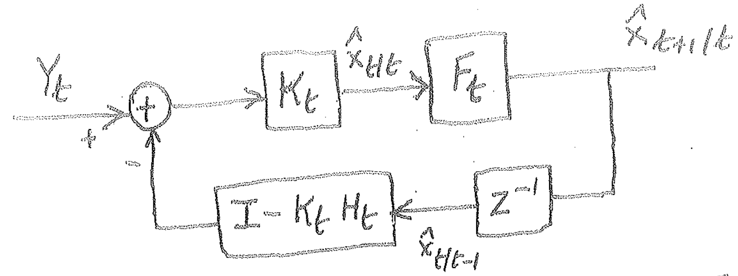
- Summarizing, (8) and (9) give an update for the error covariance matrices:

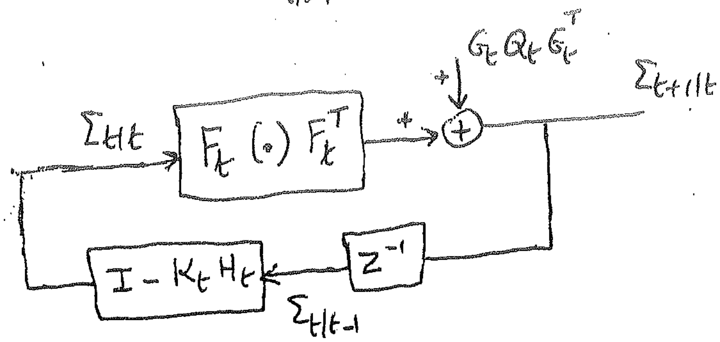
$$\boxed{\begin{aligned} \Sigma_{t+1|t} &= F_t \Sigma_{t|t} F_t^T + G_t Q_t G_t^T \quad (8) \\ \Sigma_{t|t} &= \Sigma_{t|t-1} - K_t H_t \Sigma_{t|t-1} \quad (9) \end{aligned}} \qquad t \geq 0$$

Where K_t is given in (7) in terms of $\Sigma_{t|t-1}$.

Initial conditions are $\Sigma_{0|-1} = Cov(X_0) = \Sigma_0$.

Block Diagrams:





Comments

- Covariance updates (8)(9) are data-independent.
- Recursive structure: no need to store Y_0^t
  - Measurement update (6) shows $\hat{x}_{t|t}$ is a l.c. of $\hat{x}_{t|t-1}$ & correction term $I_t = Y_t - H_t \hat{x}_{t|t-1}$
  - Time update (3) projects state estimate $\hat{x}_{t|t}$ to time $t+1$ before measurement Y_{t+1} is taken.
  - Covariance update eqns (8)(9) admit similar interpretation.
- Can eliminate $\hat{x}_{t|t}$ and $\Sigma_{t|t}$ from equations above, e.g.,

$$\hat{x}_{t+1|t} = F_t \hat{x}_{t|t-1} + F_t K_t (Y_t - H_t \hat{x}_{t|t-1})$$

$$\Sigma_{t+1|t} = F_t \Sigma_{t|t-1} F_t^T - F_t K_t H_t \Sigma_{t|t-1} F_t^T + G_t Q_t G_t^T$$

<u>Riccati eqn</u>

$$\Sigma_{t|t-1} H_t^T (H_t \Sigma_{t|t-1} H_t^T + R_t)^{-1}$$

From T. K. Moon and W. C. Stirling,
Mathematical Methods & Algorithms for Signal Processing, Prentice Hall, 2000

Example 13.2.1 In our first example of Kalman filtering, we will take a simple setup, explaining subsequently (in section 13.5) our rationale for the structure of the example. We consider simple kinematic motion in two coordinates, $(x_{1,t}, x_{2,t})$, with a motion update according to

$$x_{i,t+1} = x_{i,t} + \Delta \dot{x}_{i,t} + w_{i,t},$$

where $\dot{x}_{i,t}$ is the velocity in the ith direction and Δ represents some sampling time. We also assume that the velocity is subject to random fluctuations,

$$\dot{x}_{i,t+1} = \dot{x}_{i,t} + \dot{w}_t.$$

Stacking up the state vector as $\mathbf{x}_t = [x_{1,t}, \dot{x}_{1,t}, x_{2,t}, \dot{x}_{2,t}]^T$, we have the state equation

$$\mathbf{x}_{t+1} = \begin{bmatrix} 1 & \Delta & 0 & 0 \\ 0 & 0 & 1 & \Delta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_t + \mathbf{w}_t,$$

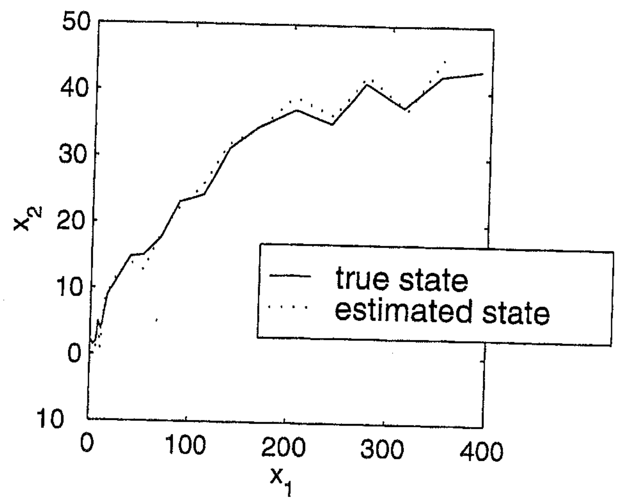


Figure 13.1: Illustration of Kalman filter

where $\mathbf{w}_t$ is zero-mean and Gaussian. For this example, the covariance is chosen to be

$$\operatorname{cov}(\mathbf{w}_t) = Q = 2I.$$

We assume, simply, that we observe the position variables in noise,

$$\mathbf{y}_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{x}_t + \boldsymbol{\nu}_t,$$

where $\boldsymbol{\nu}_t$ is zero-mean, Gaussian, and with covariance (for this example)

$$\operatorname{cov}(\boldsymbol{\nu}_t) = R = 3I.$$

The code in algorithm 13.2 illustrates the simulation of this dynamical system and the estimate of its state using a Kalman filter based upon the observation $\mathbf{y}_t$. Figure 13.1 illustrates the tracking ability of the filter, where the position variables of the true state and the estimated state are shown. □