SYSTEMTEORI - THE KALMAN FILTER

1. Examples

1.1. **7.5.** Let y be a stochastic process given by the system:

$$\begin{aligned} x(t+1) &= Ax(t) + Bv(t) \\ y(t) &= Cx(t) + Dv(t) \end{aligned}$$

where $x(0)$ and $v(t)$ satisfies the usual assumptions. Determine a Kalman filter for this case.

Note that in the normal Kalman filter framework we have the following assumptions:

1: $x(0) = x_0$
2: $Ex_0x_0^T = P_0$
3: $Ex_0 = 0$
4: $Ev(t)v^T(s) = \delta_{t,s}$
5: x_0 and $v(t)$ uncorrelated gaussian processes.

We will determine a recursive algorithm to compute the prediction $\hat{x}(t+1)$ given $\hat{x}(t)$ and $y(t)$. We will follow the same method as in chapter 9 of the compendium, and therefore we will use the same notation. We have that:

$$\begin{aligned} \hat{x}(t+1) &= E^{H_t(y)}x(t+1) = \{\text{remember that } H_t(y) = H_{t-1} \oplus [\tilde{y}(t)]\} \\ &= E^{H_{t-1}(y)}x(t+1) + E^{[\tilde{y}(t)]}x(t+1) \\ &= E^{H_{t-1}(y)}(Ax(t) + Bv(t)) + E^{\tilde{y}(t)}x(t+1) = A\hat{x}(t) + K(t)\tilde{y}(t) \end{aligned}$$

First, let us write $\tilde{y}(t)$ in a different way:

$$\begin{aligned} \tilde{y}(t) &= y(t) - E^{H_{t-1}(y)}y(t) = y(t) - E^{H_{t-1}y}(Cx(t) + Dv(t)) \\ &= y(t) - C\hat{x}(t) = C\tilde{x}(t) + Dv(t) \end{aligned}$$

Then, we can determine $K(t)$ with the Projection Theorem. Note that since $E^{[\tilde{y}(t)]}x(t+1) = K(t)\tilde{y}(t)$, we have by the Projection Theorem that

$$x(t+1) - K(t)\tilde{y}(t) \perp \tilde{y}(t),$$

and hence

$$\begin{aligned} & E\left[(x(t+1) - K(t)\tilde{y}(t))\tilde{y}(t)^T\right] = 0 \\ \Rightarrow\quad & Ex(t+1)\tilde{y}(t) = K(t)E\tilde{y}(t)\tilde{y}(t)^T \\ \Rightarrow\quad & K(t) = Ex(t+1)\tilde{y}(t)^T[E\tilde{y}(t)\tilde{y}(t)^T]^{-1} \end{aligned}$$

We need to determine the quantities $Ex(t+1)\tilde{y}(t)^T$ and $E\tilde{y}(t)\tilde{y}(t)^T$:

$$\begin{aligned} Ex(t+1)\tilde{y}(t)^T &= E(Ax(t)+Bv(t))(\tilde{x}(t)^TC^T+v(t)^TD^T) \\ &= E(A\tilde{x}(t)+A\hat{x}(t)+Bv(t))(\tilde{x}(t)^TC^T+v(t)^TD^T) \\ &= AE\tilde{x}(t)\tilde{x}(t)^TC^T+BEv(t)v(t)^TD^T \\ &= AP(t)C^T+BD^T \end{aligned}$$

where $P(t) \triangleq E\tilde{x}(t)\tilde{x}(t)^T$. Then, we have that:

$$\begin{aligned} E\tilde{y}(t)\tilde{y}(t)^T &= E(C\tilde{x}(t)+Dv(t))(\tilde{x}(t)^TC^T+v(t)^TD^T) \\ &= CE\tilde{x}(t)\tilde{x}(t)^TC^T+DEv(t)v(t)^TD^T \\ &= CP(t)C^T+DD^T \end{aligned}$$

So, we have the following:

$$K(t) = (AP(t)C^T+BD^T)(CP(t)C^T+DD^T)^{-1}$$

Finally, we need to determine a recursive equation for $P(t)$. First, let us determine the dynamic of $\tilde{x}(t)$. We have that:

$$\begin{aligned} \tilde{x}(t+1) &= x(t+1)-\hat{x}(t+1) \\ &= Ax(t)+Bv(t)-A\hat{x}(t)-K(t)(C\tilde{x}(t)+Dv(t)) \\ &= (A-K(t)C)\tilde{x}(t)+(B-K(t)D)v(t) \end{aligned}$$

Hence, the dynamic for $P(t+1)$ is:

$$\begin{aligned} P(t+1) &= E\tilde{x}(t+1)\tilde{x}^T(t+1) \\ &= E[(A-K(t)C)\tilde{x}(t)+(B-K(t)D)v(t)][\tilde{x}(t)^T(A-K(t)C)^T+v(t)^T(B-K(t)D)^T] \\ &= (A-K(t)C)P(t)(A-K(t)C)^T+(B-K(t)D)(B-K(t)D)^T, \end{aligned}$$

which after some manipulations and by replacing $K(t)$ turns out to be

$$P(t+1) = AP(t)A^T-(AP(t)C^T+BD^T)(CP(t)C^T+DD^T)^{-1}(AP(t)C^T+BD^T)^T+BB^T.$$

1.2. **7.1.** Consider the following system:

$$\begin{aligned} x(t+1) &= Ax(t)+Bv(t) \\ y(t) &= Cx(t) \end{aligned}$$

with

$$\begin{aligned} x(0) &= x_0 \\ Ev(t)v^T(s) &= \delta_{t,s}I \\ Ex_0x_0^T &= P_0 \end{aligned}$$

and x_0 and $v(t)$ uncorrelated, zero mean Gaussian variables.

a: Determine the Kalman filter for this case, and show that the usual Kalman filter converges at this one when $D \to 0$.

b: Use the previous result for the following scalar case:

$$\begin{aligned} x(t+1) &= \frac{1}{2}x(t)+v(t) \\ y(t) &= x(t) \end{aligned}$$

a: We will obtain the Kalman filter in the same way as in the previous exercise. We have that:

$$\hat{x}(t+1) = E^{H_t(y)}x(t+1) = E^{H_{t-1}(y)}x(t+1) + E^{[\tilde{y}]}x(t+1) = A\hat{x}(t) + K_t\tilde{y}(t)$$

where:

$$\tilde{y}(t) = y(t) - E^{H_{t-1}(y)}y(t) = y(t) - E^{H_{t-1}}(Cx(t)) = y(t) - C\hat{x}(t) = C\tilde{x}(t)$$

The projection theorem gives:

$$K(t) = Ex(t+1)\tilde{y}(t)^T(E\tilde{y}(t)\tilde{y}(t)^T)^{-1}$$

Besides, we have that:

$$Ex(t+1)\tilde{y}(t)^T = E\tilde{x}(t+1)\tilde{y}(t)^T + E\hat{x}(t+1)\tilde{y}(t)^T = E(A\tilde{x}(t)+Bv(t))(\tilde{x}(t)^TC^T) = AP(t)C^T$$

Then, we have that:

$$E\tilde{y}(t)\tilde{y}(t)^T = E(C\tilde{x}(t))(\tilde{x}(t)^TC^T) = CE\tilde{x}(t)\tilde{x}(t)^TC^T = CP_tC^T$$

So, we have the following:

$$K(t) = (AP(t)C^T)(CP(t)C^T)^{-1}$$

Finally, we need to determine a recursive equation for $P(t)$. First, let us determine the dynamic of $\tilde{x}(t)$. We have that:

$$\begin{aligned}\tilde{x}(t+1) &= x(t+1) - \hat{x}(t+1) = Ax(t) + Bv(t) - A\hat{x}(t) - K(t)(C\tilde{x}(t)) \\ &= (A - K(t)C)\tilde{x}(t) + Bv(t)\end{aligned}$$

Hence, the dynamic for $P(t+1)$ is:

$$\begin{aligned}P(t+1) &= E\tilde{x}(t+1)\tilde{x}^T(t+1) \\ &= (A - K(t)C)P(t)(A - K(t)C)^T + BB^T\end{aligned}$$

If we insert the value for $K(t)$ we have found previously, we get that:

$$P(t+1) = AP(t)A^T - AP(t)C^T(CP(t)C^T)^{-1}CP(t)A^T + BB^T$$

which is exactly the usual Kalman filter when $D \to 0$

b: In this case we have

$$A = \frac{1}{2}, \quad B = 1, \quad C = 1$$

By applying the previous results we get that:

$$K(t) = \frac{1}{2}P(t)P(t)^{-1} = \frac{1}{2}$$

and

$$P(t+1) = \left(\frac{1}{4}P(t) - \frac{1}{4}\right)P(t) + 1 = 1$$

So, the optimal filter is:

$$\hat{x}(t+1) = \frac{1}{2}y(t)$$

1.3. **Es. 3.** Consider two unknown but correlated constants, x_1 and x_2. We wish to determine the improvement in the knowledge of x_1 which is possible through a single noisy measurement of x_2. The vector and matrix quantities of interest are:

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad P_0 = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 \\ \sigma_{12}^2 & \sigma_2^2 \end{bmatrix}$$

and let r be the covariance of the measurement noise.

We can rewrite the problem in the following way:

$$\begin{aligned} x(t+1) &= x(t) \\ y(t) &= \begin{bmatrix} 0 & 1 \end{bmatrix} x(t) + rv(t) \end{aligned}$$

P_0 is the covariance matrix describing the uncertainty in x before the measurement. That is, σ_1^2 is the initial mean square error in knowledge of x_1, σ_2^2 is the initial mean square error in knowledge of x_2, and σ_{12} measures the corresponding cross-correlation.

After one measurement, the updated covariance matrix, $P(1)$, is given by:

$$P(1) = \begin{bmatrix} \sigma_1^2\left(\frac{\sigma_2^2(1-\rho^2)+r^2}{\sigma_2^2+r^2}\right) & \sigma_{12}^2\left(\frac{r^2}{\sigma_2^2+r^2}\right) \\ \sigma_{12}^2\left(\frac{r^2}{\sigma_2^2+r^2}\right) & \sigma_2^2\left(\frac{r^2}{\sigma_2^2+r^2}\right) \end{bmatrix}$$

where ρ is the correlation coefficient, defined as:

$$\rho = \frac{\sigma_{12}^2}{\sigma_1\sigma_2}$$

A few limiting cases are worth examining. First, in the case where the measurement is perfect, i.e. $r = 0$, the final uncertainty in the estimate of x_2, $P_{22}(1)$, is zero. Also, when $\rho = 0$, the final uncertainty in the estimate of x_1 is equal to the initial uncertainty: nothing can be learned from the measurement in this case. Finally, in the case where $\rho = \pm 1$, the final uncertainty in the estimate of x_1 is given by:

$$P_{11}(1) = \sigma_1^2\left(\frac{1}{1+\sigma_2^2/r^2}\right)$$

and the amount of information gained (i.e. the reduction in $P_{11}(1)$) depends on the ration of the initial mean square error in the knowledge of x_2 to the mean square error in the measurement of x_2. All these results are clearly very intuitive.

1.4. **Es. 4.** Consider the $n-$dimensional Brownian motion $w(t)$:

(1) $w(t)$ continuous a.s.
(2) $w(0) = 0$
(3) independent increments
(4) $w(t) - w(s) \in N(0, t-s)$.

Suppose that we measure $w(t)$ at the time steps: $t = 1, 2, ...$ with a measurement noise of variance σ^2.

a: Determine the optimal estimator of $w(t)$ and an equation for the steady state predictor.

b: Consider the one dimensional case, and let $\sigma = 1/2$. Plot a realization of the process, of the measurement and of the optimal estimation.

c: Vary σ and plot the Kalman gain together with the corresponding steady-state Kalman gain.

d: Consider the 2-dimensional case with

$$\sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1/10 \end{bmatrix}$$

Plot a realization of the process, the measurement and the optimal estimation.

a: First, we need a discrete model for the Brownian motion, and the measurement:

$$\begin{aligned} x(t+1) &= x(t) + u(t) \\ y(t) &= x(t) + \sigma v(t), \end{aligned}$$

where $x(t) = w(t)$. Hence, we have a standard formulation of the estimation problem with:

$$\begin{matrix} A = I & B = I \\ C = I & D = \sigma \end{matrix}$$

Besides, we know that the initial condition is $x_0 = 0$ and we have that:

$$Ex_0x_0^T = P_0 = 0$$

The solution is given by:

$$\begin{aligned} K(t) &= AP(t)C^T(CP(t)C^T + DD^T)^{-1} \\ &= P(t)(P(t) + \sigma^2 I)^{-1} \end{aligned}$$

$$\begin{aligned} P(t+1) &= AP(t)A^T - AP(t)C^T(CP(t)C^T + DD^t)^{-1}CP(t)A^T + BB^T \\ &= P(t) - P(t)(P(t) + \sigma^2 I)^{-1}P(t) + I \end{aligned}$$

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + K(t)(y(t) - C\hat{x}(t)) \\ &= \hat{x}(t) + K(t)(y(t) - \hat{x}(t)) \end{aligned}$$

The steady-state solution of the filter is obtained by imposing that $P(t+1) = P(t)$:

$$P(t+1) = P(t) \Rightarrow P(t)^2 = P(t) + \sigma^2 I$$

b: In the scalar case we can easily compute the steady state solution:

$$p^2 - p - \sigma^2 = 0 \Rightarrow p = \frac{1}{2} + \sqrt{\sigma^2 + 1/4}$$

The steady-state Kalman gain becomes:

$$K = \frac{p}{p + \sigma^2}$$

In Fig. 1 the plot of a realization of the process, of the measurement, and of the estimation is shown.

c: In Fig. 2, the time varying Kalman gain is plotted for different values of σ.

d: In Fig. 3 the 2-dimensional case is shown.

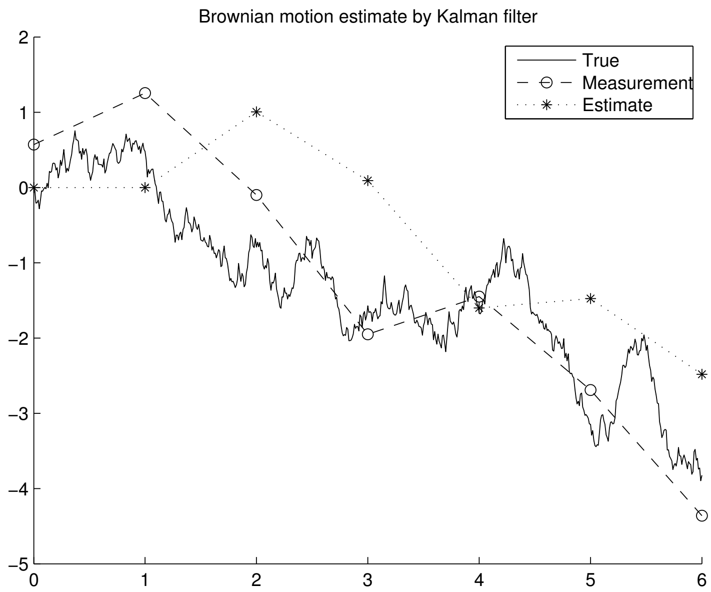

FIGURE 1. 1-dimensional Brownian motion, measurement and estimation with $\sigma = 0.5$

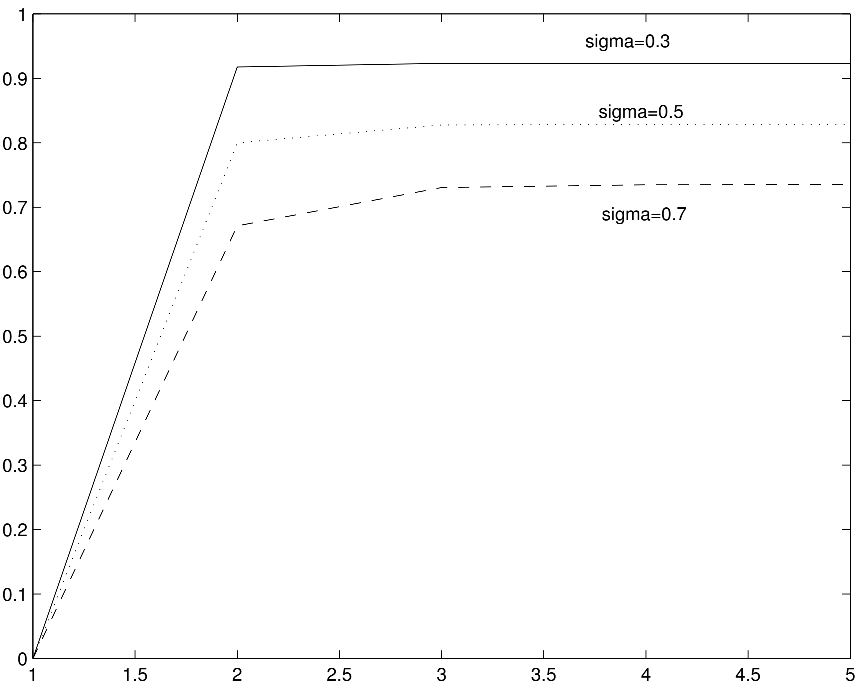

FIGURE 2. Kalman gain $K(t)$ for different values of σ

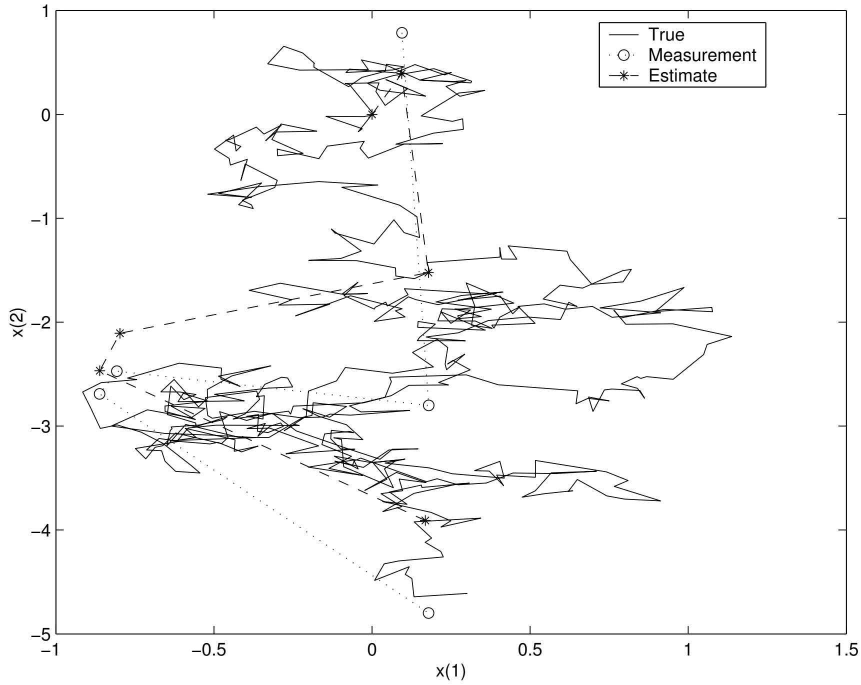

FIGURE 3. 2-dimensional Brownian motion, measurement and estimation.