Section 7
EE278: Introduction to Statistical Signal Processing (Fall 2020)

Monday, Nov 2, 2020 - 3 to 4 pm

1. Kalman filter for location tracking

Consider a truck on frictionless, straight rails [1]. Initially, the truck is stationary at location $L_0 = 0$ and moving with velocity $V_0 = 1$, but it is buffeted by random uncontrolled forces. We measure the position of the truck every $\Delta t = 1$ seconds, but these measurements are imprecise; we want to maintain a model of the truck's location L_t and its velocity V_t.

Specifically, assuming at time $t = 0$, the initial state of the truck is $L_0 = 0$, $V_0 = 1$. Between $t-1$ and t, the velocity is subject to a constant acceleration $A_{t-1} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_a^2)$. Also at time t, we take a noisy observation $Y_t = L_t + Z_t$, where $Z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_z^2)$.

Formulate this problem as a vector Kalman filter estimation problem, with $\mathbf{X}_t = [L_t, V_t]^T$.

Solution:

According to Newton's law, the dynamics can be written as

$$\begin{aligned} L_t &= L_{t-1} + V_{t-1} + \frac{1}{2} A_{t-1} \\ V_t &= V_{t-1} + A_{t-1} \end{aligned}$$

The observation can be written as

$$Y_t = L_t + Z_t$$

Define the state at time t, $\mathbf{X}_t = [L_t, V_t]^T$. Then in vector form, the above relations are

$$\begin{aligned} \mathbf{X}_t &= \begin{bmatrix} L_t \\ V_t \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{X}_{t-1} + \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} A_{t-1} \\ Y_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{X}_t + Z_t. \end{aligned}$$

Through this formulation, even though we only observe a noisy version of L_t, we can use this to estimate both L_t and V_t since they depend on each other through the A_{t-1}'s.

2. Kalman filter with fading coefficient

Consider the scalar Kalman filter specified by $X_1 \sim \mathcal{N}(\bar{X}, \sigma_{X_1}^2)$, and

$$\begin{aligned} X_n &= \alpha X_{n-1} + W_n \text{ for } n = 2, 3, ... \\ Y_n &= h X_n + Z_n \text{ for } n = 1, 2, 3, ... \end{aligned}$$

where $W_n \sim \mathcal{N}(0, \sigma_w^2)$ and $Z_n \sim \mathcal{N}(0, \sigma_z^2)$ are IID and independent each other and X_1. How do the following things change from the ones derived in class, in the presence of h?

(a) The estimate $\widehat{X}_1(y_1)$ and its mean squared error, i.e. v_1^2

(b) The estimate $\widehat{X}_2(y_1, y_2)$ and its mean squared error, i.e. v_2^2

Solution:

[1] Example taken from `https://en.wikipedia.org/wiki/Kalman_filter#Example_application`

(a) For $n = 1$, we have $X_1 \sim \mathcal{N}(\bar{X}, \sigma^2_{X_1})$ and $Y_1 = hX_1 + Z_1$. Thus,

$$\bar{Y_1} = h\bar{X}, \qquad \mathrm{Var}(Y_1) = h^2\sigma^2_{X_1} + \sigma^2_z,$$

$$\begin{aligned}
\mathrm{Cov}(X_1, Y_1) &= \mathbb{E}\left[(X_1 - \bar{X})(Y_1 - \bar{Y_1})\right] \\
&= \mathbb{E}\left[(X_1 - \bar{X})(hX_1 + Z_1 - h\bar{X})\right] \\
&= \mathbb{E}\left[h(X_1 - \bar{X})(X_1 - \bar{X})\right] + \mathbb{E}\left[(X_1 - \bar{X})Z_1\right] \\
&= h\sigma^2_{X_1} + 0
\end{aligned}$$

The estimator of X_1 given Y_1 is then

$$\begin{aligned}
\widehat{X}_1(Y_1) = \mathbb{E}[X_1 \mid Y_1 = y_1] &= \bar{X} + \frac{\mathrm{Cov}(X_1, Y_1)}{\mathrm{Var}(Y_1)}(Y_1 - \bar{Y_1}) \\
&= \bar{X} + \frac{h\sigma^2_{X_1}}{h^2\sigma^2_{X_1} + \sigma^2_z}(Y_1 - h\bar{X}) \qquad (1)
\end{aligned}$$

The mean squared error is

$$\begin{aligned}
v_1^2 = \mathrm{Var}(X_1 \mid Y_1) &= \mathrm{Var}(X_1) - \frac{\mathrm{Cov}(X_1, Y_1)^2}{\mathrm{Var}(Y_1)} \\
&= \sigma^2_{X_1} - \frac{h^2\sigma^4_{X_1}}{h^2\sigma^2_{X_1} + \sigma^2_z} \\
&= \frac{\sigma^2_{X_1}\sigma^2_z}{h^2\sigma^2_{X_1} + \sigma^2_z} \qquad (2)
\end{aligned}$$

We can write this as

$$\frac{1}{v_1^2} = \frac{1}{\sigma^2_{X_1}} + \frac{h^2}{\sigma^2_z} \qquad (3)$$

Note the differences between (1),(2),(3) and the corresponding equations derived in Lecture 13, in particular where the factor h appears.

From (3), you can see that v_1^2 is the same as in the case when the input is not scaled and the noise variance is σ^2_z/h^2 instead. This is a useful way to remember (3).

(b) With $\widehat{X}_1(y_1)$ and v_1^2 as in equations (1) and (2), we have

$$\mathbb{E}[X_2 \mid Y_1 = y_1] = \alpha\widehat{X}_1(y_1), \qquad \mathrm{Var}(X_2 \mid Y_1) = \alpha^2 v_1^2 + \sigma^2_w = u_1^2$$

Thus, by using $\mathbb{E}[X_2 \mid Y_1 = y_1]$ instead of $\bar{X}$ and $\mathrm{Var}(X_2 \mid Y_1)$ instead of $\sigma^2_{X_1}$, the estimate at $n = 2$ is

$$\widehat{X}_2(Y_1, Y_2) = \alpha\widehat{X}_1(Y_1) + \frac{hu_1^2}{h^2u_1^2 + \sigma^2_z}(Y_2 - h\alpha\widehat{X}_1(Y_1)) \qquad (4)$$

The mean squared error is

$$\begin{aligned}
v_2^2 &= v_1^2 - \frac{h^2u_1^4}{h^2u_1^2 + \sigma^2_z} \\
&= \frac{u_1^2\sigma^2_z}{h^2u_1^2 + \sigma^2_z} \qquad (5)
\end{aligned}$$

Again, this can be written as

$$\begin{aligned}\frac{1}{v_2^2} &= \frac{1}{u_1^2} + \frac{h^2}{\sigma_z^2} \\ &= \frac{1}{\alpha^2 v_1^2 + \sigma_w^2} + \frac{h^2}{\sigma_z^2} \qquad (6)\end{aligned}$$

We can generalize this to $n > 2$ as

$$u_{n-1}^2 = \alpha^2 v_{n-1}^2 + \sigma_w^2 \qquad (7)$$

$$\widehat{X}_n(Y_1, \ldots, Y_n) = \alpha \widehat{X}_{n-1} + \frac{h u_{n-1}^2}{h^2 u_{n-1}^2 + \sigma_z^2}(Y_n - h\alpha \widehat{X}_{n-1}) \qquad (8)$$

$$\frac{1}{v_n^2} = \frac{1}{\alpha^2 v_{n-1}^2 + \sigma_w^2} + \frac{h^2}{\sigma_z^2} \qquad (9)$$

Again, compare these equations with the corresponding ones derived in Lecture 13 to see how they depend on h.