A Geometric Derivation of the Scalar Kalman Filter

EECS 126 (UC Berkeley)

Spring 2019

1 Introduction

In this note, we develop an intuitive and geometric derivation of the scalar Kalman filter. Consider the following state space equations:

$$x_n = ax_{n-1} + v_n, \tag{1}$$
$$y_n = cx_n + w_n \tag{2}$$

for each positive integer n, where $(v_n)_{n=1}^\infty$ and $(w_n)_{n=1}^\infty$ are independent sources of noise. A typical scenario to keep in mind is to have a particle with position x_n moving according to the updates in (1) while measurements of the particle's position are observed as in (2). We will additionally restrict our attention to the case when $|a| < 1$. If this condition does not hold, it is possible to add a *control* term, however we will not discuss this here. Rather, **our goal is to determine** $L[x_n \mid y_1, \dots, y_n]$.

Without loss of generality, we assume $c = 1$. Indeed, if $c = 0$, then the observations are not correlated with the particle's position, so this case is uninteresting. Otherwise, if $c \neq 0$, then we can rescale (2):

$$\frac{y_n}{c} = x_n + \frac{w_n}{c}.$$

Then, we can consider $(y_n/c)_{n=1}^\infty$ to be the new observations and $(w_n/c)_{n=1}^\infty$ to be the new observation noise variables.

2 Derivation of the Scalar Kalman Filter

We begin with the key observation from [1, Theorem 8.2].

Lemma 1. *Assume that X, Y, Z are zero-mean random variables. Then:*

$$L[X \mid Y, Z] = L[X \mid Y] + L\big[X \bigm| Z - L[Z \mid Y]\big].$$

How does Lemma 1 help us? We are interested in:

$$\begin{aligned}\hat{x}_{n|n} &:= L[x_n \mid y_1, \ldots, y_n] \\ &= L[x_n \mid y_1, \ldots, y_{n-1}] + L\big[x_n \bigm| y_n - L[y_n \mid y_1, \ldots, y_{n-1}]\big]\end{aligned}$$

The first quantity in the sum is the best estimate of x_n given the observations $y_1, \ldots, y_{n-1}$, let us denote it $\hat{x}_{n|n-1}$. Additionally, we call

$$\tilde{y}_n = y_n - L[y_n \mid y_1, \ldots, y_{n-1}]$$

the **innovation** in y_n. Thus, we have:

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + k_n \tilde{y}_n \tag{3}$$

which is our first Kalman filter equation. We note that $\hat{x}_{n|n-1} = a\hat{x}_{n-1|n-1}$, so that if we are estimating online we have access to this quantity. Additionally,

$$\begin{aligned}\tilde{y}_n &= y_n - L[y_n \mid y_1, \ldots, y_{n-1}] = y_n - L[x_n + w_n \mid y_1, \ldots, y_{n-1}] \\ &= y_n - L[x_n \mid y_1, \ldots, y_{n-1}] = y_n - \hat{x}_{n|n-1}.\end{aligned}$$

Thus, we see that if we can determine the quantity k_n (referred to as the Kalman gain), we are done. To do this, we proceed geometrically as in Figure 1. How does one arrive at such a diagram? First, we place the origin 0 and x_n. This does not violate any constraints as we are simply orienting ourselves and placing an arbitrary vector. Now, we would like to draw the vector corresponding to $\hat{x}_{n|n-1}$. The only constraint given the vectors thus far is that $\hat{x}_{n|n-1} \perp (x_n - \hat{x}_{n|n-1})$ and placing $\hat{x}_{n|n-1}$ as in Figure 1 satisfies this. Now, we place the vector corresponding to $\tilde{y}_n$. We thus need $\tilde{y}_n \perp \hat{x}_{n|n-1}$, so we draw it as in Figure 1. Vector addition thus fixes the position of y_n. Additionally, we project x_n onto $\tilde{y}_n$ to get the vector $k_n\tilde{y}_n$. We are now ready to find k_n geometrically.

Note that the triangles with vertices $(\hat{x}_{n|n-1}, x_n, y_n)$ is similar to the triangle with vertices $(\hat{x}_{n|n-1}, \hat{x}_{n|n}, x_n)$, and thus

$$\frac{\|\hat{x}_{n|n} - \hat{x}_{n|n-1}\|}{\|x_n - \hat{x}_{n|n-1}\|} = \frac{\|x_n - \hat{x}_{n|n-1}\|}{\|y_n - \hat{x}_{n|n-1}\|}.$$

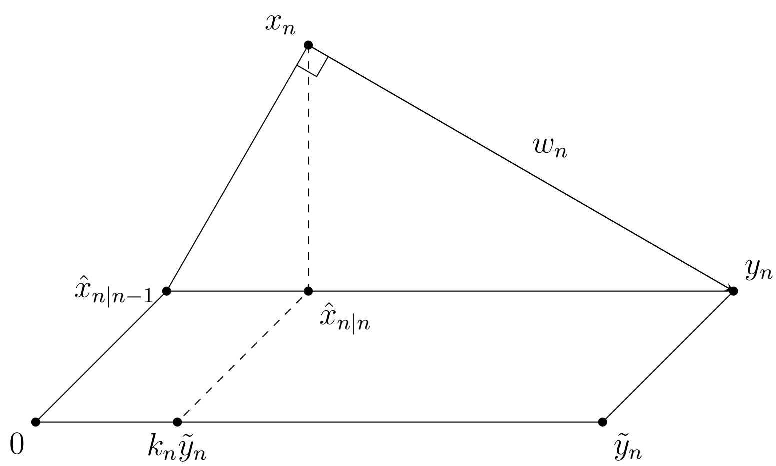

Figure 1: Geometry of the Kalman filter.

Now, since $\|\hat{x}_{n|n} - \hat{x}_{n|n-1}\| = k_n \|y_n - \hat{x}_{n|n-1}\|$, by rearranging one has

$$k_n = \frac{\|x_n - \hat{x}_{n|n-1}\|^2}{\|y_n - \hat{x}_{n|n-1}\|^2} = \frac{\sigma^2_{n|n-1}}{\sigma^2_{n|n-1} + \sigma^2_w}. \tag{4}$$

The denominator of this last equality comes from the right triangle with vertices $(\hat{x}_{n|n-1}, x_n, y_n)$. We know σ^2_w, so it remains to compute $\sigma^2_{n|n-1}$. In order to find this, we need another picture. [1] Although we went through the construction of Figure 1 in detail, we will simply give Figure 2.

Noting that we are interested in $\sigma^2_{n|n-1}$, we examine the triangle with vertices $(\hat{x}_{n|n-1}, ax_{n-1}, x_n)$. Note that by similar triangles,

$$\|ax_{n-1} - \hat{x}_{n|n-1}\| = a\|\Delta_{n-1|n-1}\|$$

and that $\|\Delta_{n|n-1}\|^2 = \|ax_{n-1} - \hat{x}_{n|n-1}\|^2 + \|v_{n-1}\|^2$, so

$$\sigma^2_{n|n-1} = a^2\sigma^2_{n-1|n-1} + \sigma^2_v. \tag{5}$$

This implies we need one final quantity: $\sigma^2_{n|n}$. Once we have this, in each iteration, we can simply pass along $\sigma^2_{n|n}$. To find this, we again examine

[1] Interestingly, it is sufficient to use one 4-D plot to draw all that we need, but this is hard (impossible?) to visualize, so we draw another 3-D plot.

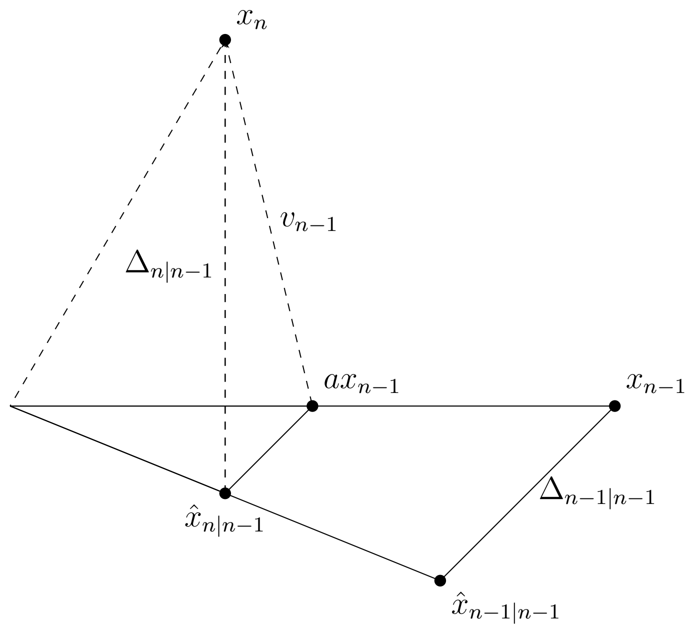

Figure 2: Geometry of the Kalman filter.

Figure 1. We note that $\sigma^2_{n|n} = \|x_n - \hat{x}_{n|n}\|^2$ and $\sigma^2_{n|n-1} = \|x_n - \hat{x}_{n|n-1}\|^2$. By the Pythagorean Theorem, we know that

$$\|x_n - \hat{x}_{n|n-1}\|^2 = \|\hat{x}_{n|n} - \hat{x}_{n|n-1}\|^2 + \|x_n - \hat{x}_{n|n}\|^2.$$

Thus,

$$\begin{aligned}\sigma^2_{n|n} = \|x_n - \hat{x}_{n|n}\|^2 &= \|x_n - \hat{x}_{n|n-1}\|^2 - \|\hat{x}_{n|n} - \hat{x}_{n|n-1}\|^2 \\ &= \|x_n - \hat{x}_{n|n-1}\|^2 \left(1 - \frac{\|\hat{x}_{n|n} - \hat{x}_{n|n-1}\|^2}{\|x_n - \hat{x}_{n|n-1}\|^2}\right) \\ &= \|x_n - \hat{x}_{n|n-1}\|^2 \left(1 - \frac{\|x_n - \hat{x}_{n|n-1}\|^2}{\|y_n - \hat{x}_{n|n-1}\|^2}\right) = \sigma^2_{n|n-1}(1 - k_n).\end{aligned}$$

We have successfully derived the scalar Kalman filter equations in the case $c = 1$. The formulas are listed here:

$$\begin{aligned}\hat{x}_{n|n} &= \hat{x}_{n|n-1} + k_n \tilde{y}_n, \\ \tilde{y}_n &= y_n - a\hat{x}_{n-1|n-1},\end{aligned}$$

$$k_n = \frac{\sigma^2_{n|n-1}}{\sigma^2_{n|n-1} + \sigma^2_w},$$
$$\sigma^2_{n|n-1} = a^2\sigma^2_{n-1|n-1} + \sigma^2_v,$$
$$\sigma^2_{n|n} = \sigma^2_{n|n-1}(1 - k_n).$$

One key observation is that the gain k_n may be computed offline! Thus, in practice, one can precompute the gain, and quickly find the estimates $\hat{x}_{n|n}$ as observations stream in.

3 Vector Case

Let us now examine the case when our state is a vector. The state space equations in this case are:

$$X_n = AX_{n-1} + V_{n-1}, \tag{6}$$
$$Y_n = CX_n + W_n, \tag{7}$$

where $(V_i)_{i=1}^{\infty}$, $(W_i)_{i=1}^{\infty}$ are orthogonal, zero-mean sources of error. The vector equations are as follows:

$$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n\tilde{Y}_n, \tag{8}$$
$$\tilde{Y}_n = Y_n - C\hat{X}_{n|n-1}, \tag{9}$$
$$K_n = \Sigma_{n|n-1}C^\mathsf{T}(C\Sigma_{n|n-1}C^\mathsf{T} + \Sigma_W)^{-1}, \tag{10}$$
$$\Sigma_{n|n-1} = A\Sigma_{n-1|n-1}A^\mathsf{T} + \Sigma_V, \tag{11}$$
$$\Sigma_{n|n} = (I - K_nC)\Sigma_{n|n-1}. \tag{12}$$

4 Conclusion

We have presented a simple derivation of the scalar Kalman filter in this note. We did not provide a proof or the update equations for the vector case in order to keep the note less cluttered. For these, please see [1, Section 8.2].

References

[1] Jean Walrand. *Probability in Electrical Engineering and Computer Science: An Application-Driven Course*. Quorum Books, 2014.