The Kalman Filter, part 1

Andreas Jakobsson

The Kalman filter

We can estimate and predict the states using the conditional expectation

$$\hat{\mathbf{x}}_{t+k|t} = E\left\{\mathbf{x}_{t+k}|\mathbf{Y}_t\right\}$$

where

$$\mathbf{Y}_t = \left[\begin{array}{ccc} \mathbf{y}_1^T & \ldots & \mathbf{y}_t^T \end{array}\right]^T$$

The Kalman filter

We can estimate and predict the states using the conditional expectation

$$\hat{\mathbf{x}}_{t+k|t} = E\left\{\mathbf{x}_{t+k}|\mathbf{Y}_t\right\}$$

where

$$\mathbf{Y}_t = \left[\begin{array}{ccc} \mathbf{y}_1^T & \ldots & \mathbf{y}_t^T \end{array}\right]^T$$

To allow us to form this estimate recursively, we use the result

$$E\{\mathbf{x}|\mathbf{y},\mathbf{z}\} = E\{\mathbf{x}|\mathbf{z}\} + C\{\mathbf{x},\mathbf{y}|\mathbf{z}\}V\{\mathbf{y}|\mathbf{z}\}^{-1}\left(\mathbf{y} - E\{\mathbf{y}|\mathbf{z}\}\right)$$

Here, think of $\mathbf{y}$ as the most recent sample, at time t, whereas $\mathbf{z}$ denotes the measurements up to $t-1$. *Note the similarity to the RLS formulation!*

One can also express the conditional variances using

$$V\{\mathbf{x}|\mathbf{y},\mathbf{z}\} = V\{\mathbf{x}|\mathbf{z}\} - C\{\mathbf{x},\mathbf{y}|\mathbf{z}\}V\{\mathbf{y}|\mathbf{z}\}^{-1}C\{\mathbf{x},\mathbf{y}|\mathbf{z}\}^T$$

The Kalman filter

Let $\mathbf{x} = \mathbf{x}_t$, $\mathbf{y} = \mathbf{y}_t$, and $\mathbf{z} = \mathbf{Y}_{t-1}$. Then,

$$E\{\mathbf{x}|\mathbf{y},\mathbf{z}\} = E\{\mathbf{x}|\mathbf{z}\} + C\{\mathbf{x},\mathbf{y}|\mathbf{z}\}V\{\mathbf{y}|\mathbf{z}\}^{-1}\left(\mathbf{y} - E\{\mathbf{y}|\mathbf{z}\}\right)$$

implies that

$$\begin{aligned}\hat{\mathbf{x}}_{t|t} &= E\left\{\mathbf{x}_t|\mathbf{y}_t,\mathbf{Y}_{t-1}\right\}\\ &= \hat{\mathbf{x}}_{t|t-1} + \mathbf{R}^{x,y}_{t|t-1}\left[\mathbf{R}^{y,y}_{t|t-1}\right]^{-1}\left(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}\right)\\ &= \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}\right)\end{aligned}$$

where

$$\begin{aligned}\mathbf{R}^{x,y}_{t|t-1} &= C\left\{\mathbf{x}_t,\mathbf{y}_t|\mathbf{Y}_{t-1}\right\}\\ \mathbf{R}^{y,y}_{t|t-1} &= V\left\{\mathbf{y}_t|\mathbf{Y}_{t-1}\right\}\\ \mathbf{K}_t &= \mathbf{R}^{x,y}_{t|t-1}\left[\mathbf{R}^{y,y}_{t|t-1}\right]^{-1}\end{aligned}$$

Looks good, but we still need to update several of the variables. We can use

$$\begin{aligned}\hat{\mathbf{x}}_{t+1|t} &= \mathbf{A}_t\hat{\mathbf{x}}_{t|t} + \mathbf{B}_t\mathbf{u}_t\\ \hat{\mathbf{y}}_{t+1|t} &= \mathbf{C}_t\hat{\mathbf{x}}_{t+1|t}\end{aligned}$$

This only leaves the updating of $\mathbf{R}^{x,y}_{t|t-1}$ and $\mathbf{R}^{y,y}_{t|t-1}$.

The Kalman filter

Using

$$V\{\mathbf{x}|\mathbf{y},\mathbf{z}\} = V\{\mathbf{x}|\mathbf{z}\} - C\{\mathbf{x},\mathbf{y}|\mathbf{z}\}V\{\mathbf{y}|\mathbf{z}\}^{-1}C\{\mathbf{x},\mathbf{y}|\mathbf{z}\}^T$$

yields

$$\begin{aligned}
\mathbf{R}^{x,x}_{t|t} &= \mathbf{R}^{x,x}_{t|t-1} - \mathbf{R}^{x,y}_{t|t-1}\left[\mathbf{R}^{y,y}_{t|t-1}\right]^{-1}\left[\mathbf{R}^{x,y}_{t|t-1}\right]^T \\
&= \mathbf{R}^{x,x}_{t|t-1} - \mathbf{K}_t\mathbf{R}^{y,y}_{t|t-1}\mathbf{K}_t^T
\end{aligned}$$

where (for a Gaussian process)

$$\begin{aligned}
\mathbf{R}^{x,x}_{t|t-1} &= V\left\{\mathbf{x}_t|\mathbf{Y}_{t-1}\right\} = V\left\{\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}\right\} = V\left\{\tilde{\mathbf{x}}_{t|t-1}\right\} \\
\mathbf{R}^{y,y}_{t|t-1} &= V\left\{\mathbf{y}_t|\mathbf{Y}_{t-1}\right\} = V\left\{\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}\right\} = V\left\{\tilde{\mathbf{y}}_{t|t-1}\right\} \\
\mathbf{R}^{x,y}_{t|t-1} &= C\left\{\mathbf{x}_t, \mathbf{y}_t|\mathbf{Y}_{t-1}\right\} = C\left\{\tilde{\mathbf{x}}_{t|t-1}, \tilde{\mathbf{y}}_{t|t-1}\right\}
\end{aligned}$$

where $\tilde{\mathbf{x}}_{t|t-1} = \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}$ and $\tilde{\mathbf{y}}_{t|t-1} = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}$.

The Kalman filter

Recall that

$$\begin{aligned}
\mathbf{x}_{t+1} &= \mathbf{A}_t\mathbf{x}_t + \mathbf{B}_t\mathbf{u}_t + \mathbf{e}_t \\
\mathbf{y}_t &= \mathbf{C}_t\mathbf{x}_t + \mathbf{w}_t
\end{aligned}$$

Then, using

$$\begin{aligned}
\hat{\mathbf{x}}_{t+1|t} &= \mathbf{A}_t\hat{\mathbf{x}}_{t|t} + \mathbf{B}_t\mathbf{u}_t \\
\hat{\mathbf{y}}_{t+1|t} &= \mathbf{C}_t\hat{\mathbf{x}}_{t+1|t}
\end{aligned}$$

yields

$$\begin{aligned}
\tilde{\mathbf{x}}_{t+1|t} &= \mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t|t-1} \\
&= \mathbf{A}_t(\mathbf{x}_t - \hat{\mathbf{x}}_{t|t}) + \mathbf{e}_t = \mathbf{A}_t\tilde{\mathbf{x}}_{t|t} + \mathbf{e}_t \\
\tilde{\mathbf{y}}_{t|t-1} &= \mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1} \\
&= \mathbf{C}_t(\mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t+1|t}) + \mathbf{w}_t
\end{aligned}$$

Thus,

$$\begin{aligned}
V\left\{\tilde{\mathbf{x}}_{t+1|t}\right\} &= \mathbf{R}^{x,x}_{t+1|t} = \mathbf{A}_t\mathbf{R}^{x,x}_{t|t}\mathbf{A}_t^T + \mathbf{R}_e \\
V\left\{\tilde{\mathbf{y}}_{t+1|t}\right\} &= \mathbf{R}^{y,y}_{t+1|t} = \mathbf{C}_t\mathbf{R}^{x,x}_{t+1|t}\mathbf{C}_t^T + \mathbf{R}_w \\
C\left\{\tilde{\mathbf{x}}_{t+1|t}, \tilde{\mathbf{y}}_{t+1|t}\right\} &= \mathbf{R}^{x,y}_{t+1|t} = \mathbf{R}^{x,x}_{t+1|t}\mathbf{C}_t^T
\end{aligned}$$

And... we are done!

The Kalman filter

The Kalman filter

$$\begin{aligned}
\hat{\mathbf{x}}_{t|t} &= \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}_t\left(\mathbf{y}_t - \hat{\mathbf{y}}_{t|t-1}\right) \\
\hat{\mathbf{x}}_{t+1|t} &= \mathbf{A}_t\hat{\mathbf{x}}_{t|t} + \mathbf{B}_t\mathbf{u}_t \\
\mathbf{K}_t &= \mathbf{R}^{x,y}_{t|t-1}\left[\mathbf{R}^{y,y}_{t|t-1}\right]^{-1} = \mathbf{R}^{x,x}_{t|t-1}\mathbf{C}_t^T\left[\mathbf{R}^{y,y}_{t|t-1}\right]^{-1}
\end{aligned}$$

with

$$\begin{aligned}
\mathbf{R}^{x,x}_{t|t} &= \mathbf{R}^{x,x}_{t|t-1} - \mathbf{K}_t\mathbf{R}^{y,y}_{t|t-1}\mathbf{K}_t^T \\
&= \mathbf{R}^{x,x}_{t|t-1} - \mathbf{K}_t\mathbf{C}_t\mathbf{R}^{x,x}_{t|t-1} \\
&= (\mathbf{I} - \mathbf{K}_t\mathbf{C}_t)\,\mathbf{R}^{x,x}_{t|t-1} \\
\mathbf{R}^{x,x}_{t+1|t} &= \mathbf{A}_t\mathbf{R}^{x,x}_{t|t}\mathbf{A}_t^T + \mathbf{R}_e \\
\mathbf{R}^{y,y}_{t+1|t} &= \mathbf{C}_t\mathbf{R}^{x,x}_{t+1|t}\mathbf{C}_t^T + \mathbf{R}_w
\end{aligned}$$

As initial conditions, one should select

$$\begin{aligned}
\hat{\mathbf{x}}_{1|0} &= E\{\mathbf{x}_1\} = \mathbf{m}_0 \\
\mathbf{R}^{x,x}_{1|0} &= V\{\mathbf{x}_1\} = \mathbf{V}_0
\end{aligned}$$

Here, $\mathbf{V}_0$ indicates your trust in $\mathbf{m}_0$; if you are confident in your estimate of $\mathbf{m}_0$, select $\mathbf{V}_0$ *small*, otherwise *large*.

For example, you can select $\mathbf{m}_0$ as the parameters you estimated using your non-recursive model. Then, set $\mathbf{V}_0$ as the variance of this model. You likely need to tune it further, but this will be a good starting point.