Kalman Filters: Derivation

Burak Bayramlı

July 10, 2011

$$
\begin{aligned}
p(z_n|z_{n-1}) &= N(z_n|Az_{n-1}, \Gamma) \\
p(x_n|z_n) &= N(x_n|Cz_n, \Sigma)
\end{aligned}
$$

The quantity, random variable, $\alpha(z_n)$ represent the joint probability of all of the data $x_1, .., x_n$ and the value of z_n. We will try to represent $\alpha(z_n)$ in terms of $\alpha(z_{n-1})$ in order to get a recursive relation.

$$
\begin{aligned}
\alpha(z_n) &= p(x_1, .., x_n, z_n) \\
&= p(x_1, .., x_n|z_n)p(z_n) \\
&= p(x_n|z_n)p(x_1, .., x_{n-1}|z_n)p(z_n) \\
&= p(x_n|z_n)p(x_1, .., x_{n-1}, z_n) \\
&= p(x_n|z_n) \int p(x_1, .., x_{n-1}, z_{n-1}, z_n) dz_{n-1} \\
&= p(x_n|z_n) \int p(x_1, .., x_{n-1}, z_n|z_{n-1})p(z_{n-1}) dz_{n-1} \\
&= p(x_n|z_n) \int p(x_1, .., x_{n-1}|z_{n-1})p(z_n|z_{n-1})p(z_{n-1}) dz_{n-1} \\
&= p(x_n|z_n) \int p(x_1, .., x_{n-1}, z_{n-1})p(z_n|z_{n-1}) dz_{n-1}
\end{aligned}
$$

Using the definition for $\alpha(z_n)$ we see that the factor after integral looks like $\alpha(z_{n-1})$. Then we can restate the equation above in terms of α.

$$
\alpha(z_n) = p(x_n|z_n) \int \alpha(z_{n-1})p(z_n|z_{n-1}) dz_{n-1}
$$

Now all of the factors above are familiar from a Kalman filter based model. $p(x_n|z_n)$ is the emission probability, $p(z_n|z_{n-1})$ is the transition probability. These distributions will be known to us before we start filtering, hence can be used for determining each α as time goes on.

Scaling

As time passes, since at each time step we are multiplying small probability values (less than 1) with eachother, we can quickly reach very small numbers very fast. In order to save $\alpha(z_n)$ from becoming too small, we can use a normalized version.

$$\hat{\alpha}(z_n) = \frac{\alpha(z_n)}{p(x_1, .., x_n)}$$

In order to work this into our equation,

$$\begin{aligned} c_n &= p(x_n|x_1, .., x_{n-1}) \\ \prod_{m=1}^{n} c_m &= p(x_1, ..., x_n) \end{aligned}$$

The last statement follows from product rule in probability theory, say for

$$p(x_1)p(x_2|x_1)p(x_3|x_2, x_1)...$$

The first two factors result in $p(x_1, x_2)$ which, in turn multiplied by third factor result in $p(x_3, x_2, x_1)$, so on. Then, we can turn the recursive statement $\alpha(z_n)$ into a scaled version

$$\begin{aligned} \alpha(z_n) &= p(x_n|z_n) \int \alpha(z_{n-1})p(z_n|z_{n-1})dz_{n-1} \\ \left(\prod_{m=1}^{n} c_m\right)\hat{\alpha}(z_n) &= p(x_n|z_n) \int \left(\prod_{m=1}^{n-1} c_m\right)\hat{\alpha}(z_{n-1})p(z_n|z_{n-1})dz_{n-1} \end{aligned}$$

Since the only difference between two c_m products on the left and right is one c_n, once canceled out, only c_n on the left handside remains.

$$c_n\hat{\alpha}(z_n) = p(x_n|z_n) \int \hat{\alpha}(z_{n-1})p(z_n|z_{n-1})dz_{n-1}$$

where

$$\hat{\alpha}(z_n) = N(z_n|\mu_n, V_n)$$

Filtering

Substituting into recursive equation

$$c_n N(z_n|\mu_n, V_n) = N(x_n|Cz_n, \Sigma) \int N(z_n|Az_{n-1}, \Gamma) N(z_{n-1}|\mu_{n-1}, V_{n-1}) dz_{n-1}$$

References

[1] C. Bishop *Pattern Recognition and Machine Learning* , 2006.