Contents

Bayesian Filters for Static Parameters in Dynamic Models

The standard setup for filtering problems is the state-observation dynamics considered previously. An alternative framework considers the case of estimating static (ie time-invariant) parameters in a dynamically observed system. For example, assume we have an observation sequence $y_{1:T} = \{y_1, y_2, \ldots, y_T\}$ from the model

$$dy = \kappa(y - \mu) + \sigma dW, \tag{1}$$

and we wish to compute online estimates of the parameters $\theta = (\kappa, \mu, \sigma)$. The relevant posterior to use here is $p(\theta|y_{1:t})$.

To work towards this we write $p(\theta|y_{1:t})$ recursively as done previously, deriving an expression $p(\theta|y_{1:t}) = k_{t-1,t}p(\theta|y_{1:t-1})$ as follows.

$$\begin{aligned}
p\left(\theta|y_{1:t}\right) &= \frac{p\left(y_{1:t}|\theta\right)p\left(\theta\right)}{p\left(y_{1:t}\right)} \\
&= \frac{p\left(y_t|y_{1:t-1},\theta\right)p\left(y_{1:t-1}|\theta\right)p\left(\theta\right)}{p\left(y_t|y_{1:t-1},\right)p\left(y_{1:t-1}|\right)} \\
&= \frac{p\left(y_t|y_{1:t-1},\theta\right)}{p\left(y_t|y_{1:t-1}\right)}\frac{p\left(y_{1:t-1}|\theta\right)p\left(\theta\right)}{p\left(y_{1:t-1}\right)} \\
&= k_{t-1,t}p\left(\theta|y_{1:t-1}\right)
\end{aligned}$$

where

$$\begin{aligned}
k_{t-1,t} &= \frac{p\left(y_t|y_{1:t-1},\theta\right)}{p\left(y_t|y_{1:t-1}\right)} \\
&= \frac{p\left(y_t|y_{1:t-1},\theta\right)}{\int p\left(y_t|y_{1:t-1},\theta\right)p\left(\theta|y_{1:t-1}\right)d\theta}
\end{aligned}$$

is the required *transition kernal*. In this case the computational strategy is clearer after substituting $k_{t-1,t}$ back into the original expression to get

$$\begin{aligned}
p\left(\theta|y_{1:t}\right) &= k_{t-1,t}p\left(\theta|y_{1:t-1}\right) \\
&= \frac{p\left(y_t|y_{1:t-1},\theta\right)p\left(\theta|y_{1:t-1}\right)}{\int p\left(y_t|y_{1:t-1},\theta\right)p\left(\theta|y_{1:t-1}\right)d\theta}
\end{aligned}$$

To solve using a particle approach we again approximate the posterior $p\left(\theta|y_{1:t}\right)$ as a sequence of N points θ^i_t and weights $\pi^i_t = p\left(\theta^i|y_{1:t}\right)$. To get there insert the particle approximation into the expression for $p\left(\theta|y_{1:t}\right)$ above, and simplify using the Markov property assumed in our model

$$\begin{aligned}
p\left(\theta|y_{1:t}\right) &= \frac{p\left(y_t|y_{1:t-1},\theta\right)p\left(\theta|y_{1:t-1}\right)}{\int p\left(y_t|y_{1:t-1},\theta\right)p\left(\theta|y_{1:t-1}\right)d\theta} \\
&= \frac{p\left(y_t|y_{t-1},\theta\right)p\left(\theta|y_{t-1}\right)}{\int p\left(y_t|y_{t-1},\theta\right)p\left(\theta|y_{t-1}\right)d\theta}
\end{aligned}$$

and insert the particle approximation

$$\begin{aligned}
p(\theta^i|y_t) &\sim \frac{p\left(y_t|y_{t-1},\theta^i\right)p\left(\theta^i|y_{t-1}\right)}{\sum p\left(y_t|y_{t-1},\theta^i\right)p\left(\theta|y_{t-1}\right)} \\
&= \frac{p\left(y_t|y_{t-1},\theta^i\right)\pi^i_{t-1}}{\sum p\left(y_t|y_{t-1},\theta^i\right)\pi^i_{t-1}} \\
&= \frac{\widehat{\pi}^i_{t-1}}{\sum \widehat{\pi}^i_{t-1}}
\end{aligned}$$

where

$$\widehat{\pi}_t^i = \pi_{t-1}^i p\left(y_t | y_{t-1}, \theta^i\right) \tag{2}$$

The process here is almost identical then to our previous examples. After initialising a set of N particles $\left\{\theta^i, \pi^i\right\}_{i=1}^N$ at time $t = 0$, we update our weights by computing the next $\widehat{\pi}^i$ for each i using the above formula. The weights are then normalised to obtain the estimate of the posterior. The only difference here is that the particle points θ^i are not passed through a dynamic system like our model for states previously (because they are time-invariant). However, it may be the case that we modify them over time, but that is only a computational strategy and does not reflect any dynamics inherent to the model.