5.4 Belief state = set of samples (particle filtering)

The basic idea behind particle filtering\(^6\) is to approximate the belief state by a set of weighted particles or samples:

\[
P(X_t | y_{1:t}) \approx \sum_{i=1}^{N_s} w^i_t \delta(X_t, X^i_t)
\]

(In this section, \(X^i_t\) means the i’th sample of \(X_t\), and \(X_{t,i}\) means the i’th component of \(X_t\).) This is a non-parametric approach, and hence can handle non-linearities, multi-modal distributions, etc. The advantage over discretization is that the method is adaptive, placing more particles (corresponding to a finer discretization) in places where the probability density is higher.

Given a prior of this form, we can compute the posterior using importance sampling. In importance sampling, we assume the target distribution, \(\pi(x)\), is hard to sample from; instead, we sample from a proposal or importance distribution \(q(x)\), and weight the sample according to \(w^i \propto \pi(x)/q(x)\). (After we have finished sampling, we can normalize all the weights so \(\sum w^i = 1\).) We can use this to sample paths with weights

\[
w^i_t \propto \frac{P(x^i_{1:t}|y_{1:t})}{q(x^i_{1:t}|y_{1:t})}
\]

The probability of a sample path, \(P(x^i_{1:t}|y_{1:t})\), can be computed recursively using Bayes rule. Typically we will want the proposal distribution to be recursive also, i.e., \(q(x_{1:t}|y_{1:t}) = q(x_t|x_{1:t-1}, y_{1:t})q(x_{1:t-1}|y_{1:t-1})\). In this case we have

\[
w^i_t \propto \frac{P(y_t|x^i_t)P(x^i_t|x_{1:t-1}^i)P(x_{1:t-1}^i|y_{1:t-1})}{q(x^i_t|x_{1:t-1}^i, y_{1:t})q(x_{1:t-1}^i|y_{1:t-1})} = \frac{P(y_t|x^i_t)P(x^i_t|x_{1:t-1}^i)}{q(x^i_t|x_{1:t-1}^i, y_{1:t})}w^i_{t-1}
\]

\[\begin{align*}
\text{def} & \quad \bar{w}^i_t = \tilde{w}^i_t \times w^i_{t-1}
\end{align*}\]

where we have defined \(\bar{w}^i_t\) to be the incremental weight.

For filtering, we usually only care about the posterior marginal \(P(X_t|y_{1:t})\), as opposed to the full posterior \(P(X_{1:t}|y_{1:t})\). Hence we use the following proposal: \(q(x_t|x_{1:t-1}, y_{1:t}) = q(x_t|x_{1:t-1}^i, y_t)\). This means we only need to

\(^6\)Particle filtering is also known as sequential Monte Carlo, sequential importance sampling with resampling (SISR), the bootstrap filter, the condensation algorithm, survival of the fittest, etc.