

# 1 Hidden Markov Models

A **Hidden Markov Model** (HMM) is defined by an underlying system that is a Markov process with states that cannot be observed, but gives outputs that are dependent on the hidden state of the system. If we let random variable  $X_i$  represent the  $i$ -th state of the system, and  $Y_i$  represent the  $i$ -th output of the system, we have:

$$X_1 \sim \mu(x_1) \text{ and } X_n | (X_{n-1} = x_{n-1}) \sim f(x_n | x_{n-1}) \quad (1)$$

$$Y_n | (X_n = x_n) \sim g(y_n | x_n) \quad (2)$$

where  $X_1$  is defined by some initial conditions  $\mu(x_1)$ .

# 2 Particle Filters

Basically, the big idea is to use our observations of the system to determine the hidden state of the system. That is to say, we want to figure out  $p(x_n | y_{1:n})$ . Recognizing that this is a marginal is useful because then we can just say:

$$p(x_n | y_{1:n}) = \int p(x_{1:n} | y_{1:n}) dx_{1:n-1} \quad (3)$$

Setting this as a marginal of  $p(x_{1:n} | y_{1:n})$  is helpful because we can refactor this conditional distribution to see how it comes from our prior distributions of the hidden state and observations.

$$\begin{aligned} p(x_{1:n} | y_{1:n}) &= \frac{p(x_{1:n}, y_{1:n})}{p(y_{1:n})} \\ p(x_{1:n}, y_{1:n}) &= p(x_{1:n})p(y_{1:n} | x_{1:n}) \\ p(x_{1:n}) &= \mu(x_1) \prod_{i=2}^n f(x_i | x_{i-1}) \\ p(y_{1:n} | x_{1:n}) &= \prod_{i=1}^n g(y_i | x_i) \\ p(y_{1:n}) &= \int p(x_{1:n}, y_{1:n}) dx_{1:n} \\ p(x_n | y_{1:n}) &= \int \frac{\mu(x_1) \prod_{i=2}^n f(x_i | x_{i-1}) \prod_{i=1}^n g(y_i | x_i)}{\int \mu(x_1) \prod_{i=2}^n f(x_i | x_{i-1}) \prod_{i=1}^n g(y_i | x_i) dx_{1:n-1}} dx_{1:n-1} \end{aligned} \quad (4)$$

This looks really messy, but it's helpful to see that this is just integration of functions we already know  $f$  and  $g$ . In other words, we already have all the information we need to arrive at the answer.

The issue now is that this is sometimes intractable. If functions are nice, like guassians or linear distributions, this is integrable and we can arrive at a solution analytically. However, this is not always the case, which is why we need to derive an answer from numeric approximations.

# 3 Kalman Filter

So let's say you're really lucky and your initial distribution is something nice. This allows us to get a solution analytically through a method called the Kalman Filter. We can reformulate our goal as:

$$\text{Prediction step: } p(x_n | y_{1:n}) = \int f(x_n | x_{n-1}) p(x_{n-1} | y_{1:n-1}) dx_{n-1} \quad (5)$$

We assume that  $p(x_{n-1}|y_{1:n-1})$  is known due to recursion.

$$\text{Update step: } p(x_n|y_{1:n}) \propto g(y_n|x_n)p(x_n|y_{1:n-1}) \quad (6)$$

## 4 Sequential Monte Carlo Methods

So what if our distribution isn't nice? Rekt.

## 5 SIS

## 6 Bootstrap (SIR)

We resample to deal with the degeneracy problem.

Except that there's a whole new problem called the sampling impoverishment problem. There are methods that address both of these issues but they require fancy sounding things like the "Epanechnikov Kernel".