12. HMMs with uncountable state spaces

LA. Consider a $\mathbb{R}^n$-valued discrete time sequence $S=(S_1,\dots,S_n,\dots)$ where $S_i \in \mathbb{R}^n$. Let

$$p(S_1,\dots,S_n) = p(S_1)\,p(S_2|S_1)\dots p(S_n|S_{n-1})$$

Here $p(S_1)$, $p(S_i|S_{i-1})$ are probability densities. The sequence of observations $X=(X_1,\dots,X_n)$ (generated by S) is assumed to be conditional independent

$$p(X|S) = \prod_i p(x_i|S_i),$$

where $x_i \in \mathbb{R}^m$ and $p(x_i|S_i)$ are probability densities

Example 1 (mobile robot)

- $S_i \in \mathbb{R}^2$ – robot position
- $X_i \in \mathbb{R}^m$ – measurments like ultrasound, camera, ...
- $p(x_i|S_i)$ – p.d. for measurments given the position
- $p(S_i|S_{i-1};u)$ – p.d. for new pos. given old pos. and motor controls u

Remark: For the sake of simplicity we will consider homogeneous models in the following

Suppose our task is to find $p(S_n|X_{1:n})$ where $X_{1:n}$ denotes the sequence of observations $x_1,\dots,x_n$ assuming that we know S_1.

B. Kalman Filter

Notation: $\mathcal{N}(x;\mu,A)$ denotes a multivariate Gaussian distribution for $x \in \mathbb{R}^n$ with mean μ and covariance A

$$\mathcal{N}(x;\mu,A) = \frac{1}{\sqrt{(2\pi)^n \det A}} \exp\left[-\frac{1}{2}\langle x-\mu, A^{-1}(x-\mu)\rangle\right]$$

Observation

- product of two Gaussians

$$\mathcal{N}(x;\mu,A)\cdot\mathcal{N}(x;\nu,B) \sim \mathcal{N}(x;\xi,C)$$

where $C = (A^{-1}+B^{-1})^{-1}$ and $\xi = C(A^{-1}\mu + B^{-1}\nu)$

- convolution of two Gaussians

$$\int_{\mathbb{R}^n} \mathcal{N}(x;\mu,A)\,\mathcal{N}(y-x;\nu,B)\,dx = \mathcal{N}(y;\xi,C)$$

where $C = A+B$, $\xi = \mu+\nu$

Assume now, the following holds for the HMM

$$p(s_1) = \mathcal{N}(s_1;\mu_1,Q)$$

$$p(s_i|s_{i-1}) = \mathcal{N}(s_i;As_{i-1},Q)$$

$$p(x_i|s_i) = \mathcal{N}(x_i;Hs_i,R)$$

where $s_i \in \mathbb{R}^n$, $x_i \in \mathbb{R}^m$, $A:\mathbb{R}^n \to \mathbb{R}^n$, $H:\mathbb{R}^n\to\mathbb{R}^m$ are linear mapings and Q, R covariance matrices.

It follows that $p(s_n|x_{1:n})$ is a Gaussian and can be calculated recursively

$$p(s_n|x_{1:n}) \sim p(x_n|s_n)\int_{\mathbb{R}^n} p(s_n|s_{n-1})\,p(s_{n-1}|x_{1:n-1})\,ds_{n-1}$$

Let us denote

$$p(s_{n-1} | x_{1:n-1}) = \mathcal{N}(s_{n-1}; \mu_{n-1}, D_{n-1})$$

$$p(s_n | x_{1:n-1}) = \mathcal{N}(s_n; \tilde{\mu}_n, \tilde{D}_n)$$

$$p(s_n | x_{1:n}) = \mathcal{N}(s_n; \mu_n, D_n)$$

Then μ_n, D_n can be computed as follows

$$\tilde{\mu}_n = A \mu_{n-1} \qquad \tilde{D}_n = Q + A D_{n-1} A^T$$

$$\mu_n = \tilde{\mu}_n + K_n (x_n - H \tilde{\mu}_n) \qquad D_n = \tilde{D}_n - K_n H \tilde{D}_n$$

where

$$K_n = \tilde{D}_n H^T S_n^{-1}, \qquad S_n = H \tilde{D}_n H^T + R$$

Remark: A similar „filter" can be derived for the calculation of $p(s_n | x_{1:n+l})$

LC. Particle filter

In the general case there is no closed form representation of $p(s_n | x_{1:n})$. We may use sampling to estimate it:

1. Generate an i.i.d sample s_1^l, $l = 1, \dots, L$ using $p(s_1 | x_1) \sim p(s_1) p(x_1 | s_1)$

2. Given a sample s_{i-1}^l $l = 1, \dots, L$ generated by $p(s_i | x_{1:i})$, ~~generate~~ sample s_i^l as follows

$$s_i^l \sim p(x_i | s_i) p(s_i | s_{i-1} = s_{i-1}^l)$$

The finally obtained sample S_n^l, $l = 1, \dots, L$ estimates $p(S_n | X_{1:n})$ and can be used to estimate the expectation of a random variable f:

$$\int_{\mathbb{R}^n} f(S_n)\, p(S_n | X_{1:n})\, dS_n \approx \frac{1}{L} \sum_{l=1}^{L} f(S_n^l)$$