Exercise 11

Tuesday 1. December 2020 12.08

Question 3:

Recall that Bayesian filter equations are given as

- Prediction: (1) $p(x_n \mid y_1, \dots, y_{n-1}) = \int p(x_n \mid x_{n-1})\, p(x_{n-1} \mid y_1, \dots y_{n-1})\, dx_{n-1}$

- Update:

$$p(x_n \mid y_1, \dots, y_n) = \frac{p(y_n \mid x_n)\, p(x_n \mid y_1, \dots, y_{n-1})}{\int p(y_n \mid x_n)\, p(x_n \mid y_1, \dots, y_{n-1})\, dx_n} \quad \dots (2)$$

---

*1 Conditional probability.

→ probability of event A given that (knowing) B happend.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0.$$

$$P(A \cap B) =: P(A, B) = P(A \mid B)\, P(B)$$

$$P(A, B, C) = \frac{P(A, B, C)}{P(B, C)}\, P(B, C)$$

$$= P(A \mid B, C)\, \underbrace{P(B, C)}$$

$$= P(A \mid B, C)\, P(B \mid C)\, P(C)$$

product law

$$= P(C \mid A, B)\, P(A \mid B)\, P(C)$$

Changing order so on

*2. Marginal probability / Marginalization.

X_1 → the random var. corresponds to the event that number of dice shown up, for 1st dice

X_2 → —"— , for 2nd dice

$P(X_1 = x_1, X_2 = x_2)$

$X_1 \backslash X_2$	1	2	3	4	5	6	$P(X_1 = x_1)$
1	P_{11}	P_{12}	P_{13}	P_{14}	P_{15}	P_{16}	$P_{1\bullet}$
2	P_{21}	-	-	·	·	·	$P_{2\bullet}$
3	P_{31}	·	·	·	·		$P_{3\bullet}$

4	p_{41}	.	.	.	.	.	$p_{4\cdot}$
5	p_{51}	.	.	.	.	.	$p_{5\cdot}$
6	p_{61}	.	.	.	.	.	$p_{6\cdot}$
$P(X_2=x_2)$	$p_{\cdot 1}$	$p_{\cdot 2}$	$p_{\cdot 3}$	$p_{\cdot 4}$	$p_{\cdot 5}$	$p_{\cdot 6}$	1

$$P(X_1=x_1) = \sum_{i=1}^{6} P(X_1=x_1,\ X_2=i)$$

#3. Monte Carlo Integration ?

Suppose you were given some probability distribution π and you can get samples from π.

meaning: You can generate N samples $\{x_i\}$ $i=1,\dots,N$ such that x_i is identically independent distributed according to π.

$$\pi(f) := \int f(x)\ \pi(dx)$$

$\pi(dx)$ is actually probability of x in the interval of $(x, x+dx)$

probability density.

$$\pi(dx) = p(x)\ dx$$

MC integration

$$\pi^N(f) := \frac{1}{N}\sum_{i=1}^{N} f(x_i), \quad x_i \sim \pi$$

it can be shown that:

as $N \to \infty$ then $\pi(f) - \pi^N(f) \to 0$

in some sense (metric).

---

a). x^i_{n-1}, $i=1,\dots,N$, $\{x^i_{n-1}\} \sim p(x_{n-1} \mid y_1,\dots y_{n-1})$

using $\pi(dx_{n-1}) = \underset{\text{density}}{p}(x_{n-1} \mid y_1,\dots,y_{n-1})\ dx_{n-1}$

then MC integration π^N given by:

N

$$\pi^N(\phi) = \frac{1}{N}\sum_{i=1} \phi(x^i_{n-1}) \,//$$

and as before, as $N \to \infty$, $\pi^N(\phi) \to \pi(\phi)$

where $\pi(\phi) = \int \phi(x_{n-1})\, p(x_{n-1} \mid y_1, \dots, y_{n-1})\, dx_{n-1}$ (3).

- $\phi(x) = x$, then $\pi(\phi) = \mathbb{E}[x]$
- $\phi(x) = (x - \mathbb{E}[x])^2$, $\pi(\phi) = \mathbb{E}[(x - \mathbb{E}(x))^2]$ //

b). Given $\{x^i_{n-1}\} \sim p(x_{n-1} \mid y_1, \dots, y_{n-1})\, dx_{n-1}$

Then since we know the transition probability density from $\underline{x_{n-1}}$ to $\underline{x_n}$; we know.

$$\underbrace{p(x_n \mid x_{n-1})}. \to x_n = \underbrace{f(x_{n-1})} + q_n$$

you can have non Gaussian noise here

if you have $q_n \sim \mathcal{N}(0, Q)$

then we know that $q_n = x_n - f(x_{n-1})$

hence $X_n \sim \mathcal{N}(f(x_{n-1}), Q)$.

$$p(x_n \mid \underbrace{y_1, \dots, y_{n-1}}) = \int \underbrace{p(x_n \mid x_{n-1})\, p(x_{n-1} \mid y_1 \dots y_{n-1})}_{I}\, dx_{n-1}$$

here we do a marginalization.

$$I = p(\underset{A}{x_n} \mid \underset{B}{x_{n-1}})\, p(x_{n-1} \mid \underbrace{y_1, \dots, y_{n-1}}_{C})$$

$$p(A \mid B)\, p(B \mid C)$$

$$= p(A \mid B, C)\, p(B \mid C)$$

$$= p(A, B \mid C)$$

$$P(A|C) = \sum_b P(A,B|C)$$

Therefore, simulating the dynamics from x_{n-1}^i will result in samples x_n^i which approximate, or

$$x_n^i \sim \left(\int p(x_n | x_{n-1}) \, p(x_{n-1} | y_1, \ldots, y_n) \, dx_{n-1} \right) dx_n$$

$$p(x_n | y_1, \ldots, y_{n-1}) \, dx_n$$

$$x_n^i \sim \tilde{\pi}(dx_n)$$

hence the MC integration w.r.t $\tilde{\pi}$, is given by

$$\tilde{\pi}^N(\phi) = \frac{1}{N} \sum_{i=1}^{N} \phi(x_n^i) \quad \ldots\ldots (4)$$

where $$\tilde{\pi}(\phi) = \int \phi(x_n) \, p(x_n | y_1, \ldots, y_{n-1}) \, dx_n \quad \ldots (5)$$

9. If $\tilde{w}_n^i = p(y_n | x_n^i)$, then using $\tilde{w}_n^i$ inside the monte carlo integration (4) we have

$$\left(\frac{1}{N} \sum_{i=} \tilde{w}_n^i \, \phi(x_n^i) \right) \approx \int p(y_n | x_n) \, \phi(x_n) \, p(x_n | y_1, \ldots, y_{n-1}) \, dx_n \quad \ldots (6)$$

Similarly if we choose $\phi = 1$, then we get

$$\frac{1}{N} \sum_{i=1}^{N} \tilde{w}_n^i \approx \int p(y_n | x_n) \, p(x_n | y_1, \ldots, y_{n-1}) \, dx_n \quad \ldots (7)$$

Then if we divide (6) over (7) we get

$$\frac{1}{\ } \sum^{N} \tilde{w}_n^i \, \phi(x_n^i)$$

$$\frac{\frac{1}{N}\sum_{i=1}^{N} \tilde{w}_n^i \phi(x_n^i)}{\frac{1}{N}\sum_{i=1}^{N} \tilde{w}_N^i} = \underbrace{\sum_{i=1}^{N} w_n^i \phi(x_n^i)}, \quad \hat{\pi}^{(N)}(\phi)$$

$$w_n^i = \left(\frac{\tilde{w}_n^i}{\sum_{i=1}^{N} \tilde{w}_n^i} \right)$$

$\hat{\pi}^N(\phi)$ is Monte carlo approximation of $\hat{\pi}(\phi)$ where $\hat{\pi}$ is

$$\hat{\pi}(dx_n) = \frac{p(y_n \mid x_n)\, p(x_n \mid y_1, \ldots, y_{n-1})\, dx_n}{\int p(y_n \mid x_n)\, p(x_n \mid y_1, \ldots y_{n-1})\, dx_n}$$

d). $$\hat{\pi}(\phi) = \int \phi(x_n) \frac{p(y_n \mid x_n)\, p(x_n \mid y_1, \ldots, y_{n-1})\, dx_n}{\int p(y_n \mid x_n)\, p(x_n \mid y_1, \ldots, y_{n-1})\, dx_n}$$

we know that

$$p(x_n \mid y_1, \ldots, y_{n-1}, y_n)$$

$$= \frac{p(y_n \mid x_n)\, p(x_n \mid y_1, \ldots y_{n-1})}{\int p(y_n \mid x_n)\, p(x_n \mid y_1, \ldots y_{n-1})\, dx_{n-1}}$$

then $\hat{\pi}(dx_n) = p(x_n \mid y_1, \ldots, y_{n-1}, y_n)\, dx_n$.

aposterior distribution.

then since $\hat{\pi}^N$ is MC approximation of $\hat{\pi}$, then

as $N \to \infty$,

$$\hat{\pi}^N(\phi) \to \hat{\pi}(\phi).$$

after resampling

$$\frac{1}{N}\sum^{N} \phi(x_n^i) \to \int \phi(x_n)\, p(x_n \mid y_1, \ldots y_n)\, dx_n.$$

$$p(x_n \mid y_1, \ldots\ldots, y_n).$$