Topics:

- Kalman Filter
- Particle Filter
- Rejection Sampling
- Importance Sampling

Lectures 13, 14, beginning of 15

Friday, November 3, 2023

Discrete states review:

Filtering - Bayes' filter

Smoothing - Sum-product / forward-backward on HMM

Most-likely state sequence - Viterbi

Continuous states (today):

Filtering: estimate $P(s_t \mid o_{0:t})$

- Kalman Filter when transition and sensor models are linear with Gaussian noise. for exact inference.
- EKF, Unscented, etc. when not linear-Gaussian as approximate inference

Sampling from Continuous Distributions

- Rejection Sampling

  $x \sim Q$

  $z \sim \text{Uniform}[0, kQ(x)]$

  if $z \le P(x)$: accept. else reject.

  ⇒ samples approximate P.

- Importance Sampling

  $x \sim$ sample from Q.

  weight sample by $\frac{P(x)}{Q(x)}$

  ↳ statistics using weighted samples approximately statistics of P.

- Particle filter

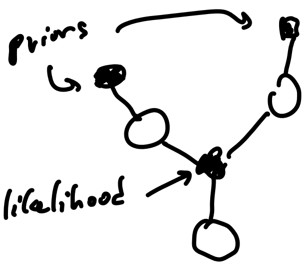
Kalman Filter

x: belief

y: evidence

$$\propto \underbrace{p(x)}_{\text{prior}} \cdot \underbrace{p(y \mid x)}_{\text{likelihood}} = \underbrace{p(x \mid y)}_{\text{posterior}}$$

"conjugacy": When prior and likelihood are Gaussian, posterior is also Gaussian.



Bayes' Net as a factor graph.

On a Bayes' network with linear-Gaussian distributions, the joint is multivariate Gaussian and conditioned on evidence, the cond. distr. is Gaussian.

Transition model $P(\underline{x}_t \mid \underline{x}_{t-1})$ same at each discrete time t and is linear - Gaussian

Sensor model $P(\underline{y}_t \mid \underline{x}_t)$ - - - - - - - - - - - -

Prior $P(\underline{x}_0)$

$\underline{x}, \underline{y}$, are vector -valued

Process Step

Probablistic view.

$$P(\underline{x}_t \mid \underline{y}_{0:t-1}) = \int P(\underline{x}_t \mid \underline{x}_{t-1}) \cdot P(\underline{x}_{t-1} \mid \underline{y}_{0:t-1})\, d\underline{x}_{t-1}$$

Since all Gaussian, can do linear algebra instead of calculus:

$$\boxed{P(\underline{x}_t \mid \underline{y}_{0:t-1}) = N(\underbrace{A x_{t-1|t-1}}_{\triangleq x_{t|t-1}}, \underbrace{A Q_{t-1|t-1} A^T + W}_{\triangleq Q_{t|t-1}})}$$

Transition Update

$$\underline{x}_t = A \underline{x}_{t-1} + \underline{w},$$

$$\underline{w} \sim N(0, W)$$

independent of all r.v.s

Measurement Step

$$P(\underline{x}_t \mid \underline{y}_{0:t}) \propto P(\underline{y}_t \mid \underline{x}_t) \cdot P(\underline{x}_t \mid \underline{y}_{0:t-1})$$

$$= \boxed{N\left(\underline{x}_{t|t-1} + K_t\left[y_t - H\underline{x}_{t|t-1}\right], \; \} \triangleq Q_{t|t} \quad Q_{t|t-1} - K_t H Q_{t|t-1} \; \} \triangleq x_{t|t}\right)}$$

Observation Update

$$\underline{y}_t = H\underline{x}_t + \underline{v},$$

$$\underline{v} \sim N(0, R)$$

independent of all r.v.s

where

K_t is 'Kalman gain'; $K_t = Q_{t|t-1} H^T R^{-1}$ ≠ end slide 18 for Kalman gain in HW

Intuitively, it weights how much to consider y_t over $x_{t|t-1}$.

Notice $\underbrace{Q_{t|t-1} \text{ ">" } Q_{t-1|t-1}}_{\text{transition}}$, $\underbrace{Q_{t|t} \text{ "<" } Q_{t|t-1}}_{\text{observation}}$.

Particle Filter

- Approximate inference
- The marginalize step as seen in Kalman Filter can quickly get cumbersome when dynamics aren't linear-Gaussian.

Ideas

- use samples as approximate representation of current state distribution.
- resample to focus on high-probability regions of state space.

Alg

- Init N particles from π.
- Update:

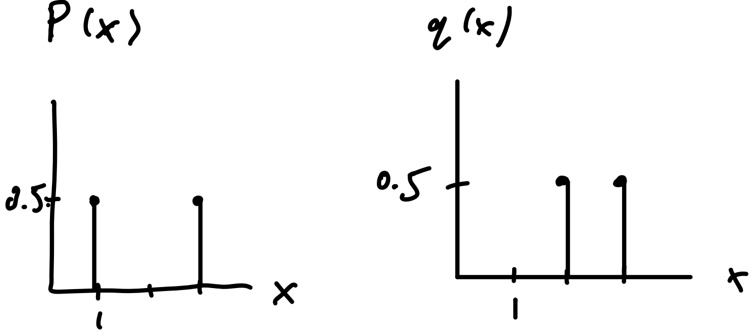
  sample $x_{t+1}^{[i]} \sim T(x_t^{[i]})$ for $i=1$ to N

  $w^{[i]} = P(y_{t+1} \mid x_{t+1}^{[i]})$ for $i=1$ to N

  Generate x_{t+1} by sampling N elts from W.

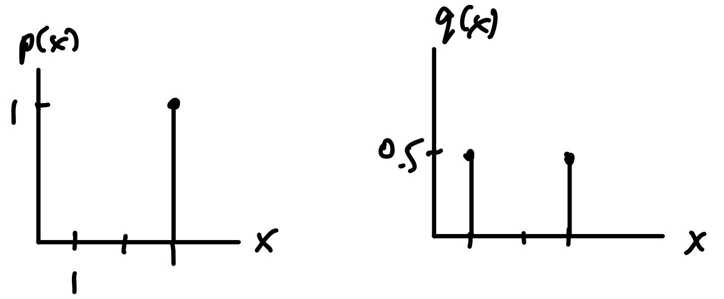
Answers

a)



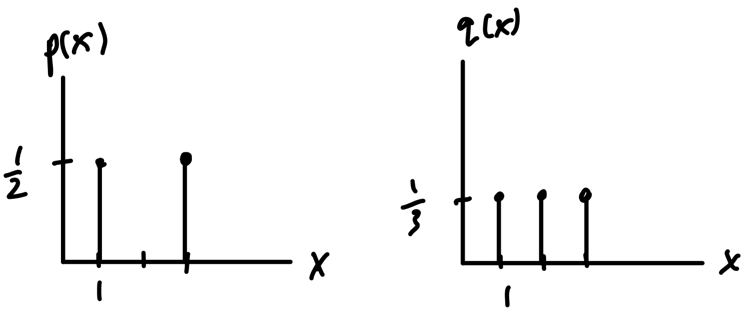
$D = \infty$

b)



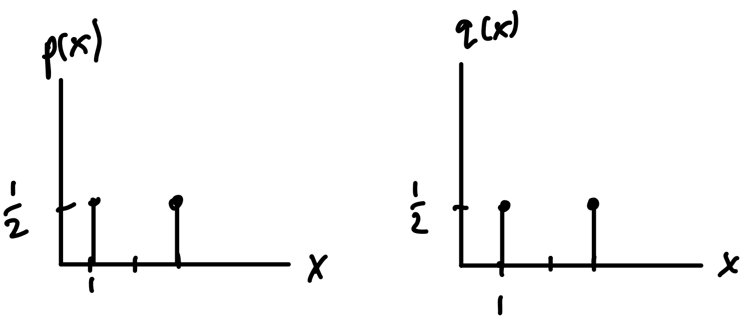
$D = 1$

c)



$D = \log_2 3 - 1$

d)



$D = 0$