Topics for the practice session on L3

Basic information

One purpose with the flipping the classroom model used in this cource is to enable us to dedicate our time in class to active learning. We call these sessions our practice sessions. During these sessions we intend to focus on problem solving, peer-instruction, multi choice questions and discussions on the fundamental concepts that we study. **These classes are for you** – take the opportunity to ask questions and discuss whatever elementary or advanced aspect of this course that you are interested in at the moment.

Today we focus on the material in **lecture** 3. To boost our discussions I have prepared a set of problems that we can look at. Though we do not have to cover all of them, I will try to limit the time that we spend on each of them to roughly 15 minutes.

Problem 1 – factorizations and Bayesian networks

In this problem we will learn more about how to factorize densities in general, and how we can factorize densities given a Bayesian network description.

To make your discussions less abstract, we encourage you to consider an example. For part a)–c) you can, for instance, assume that the variables a, b and c represent cold, high fever and visit a doctor, respectively, such that, e.g., $a = 1$ means you have a cold and $a = 0$ means that you do not have a cold.

a) Use the definition of a conditional density (the product rule) to factorise $p(a, b)$. (There are two possible solutions.)

$P(a|b)P(b)$, $P(b|a)P(a)$

b) Use the product rule to factorise $p(a, b, c)$ completely (into three factors). There are six possible solutions, but you do not have to list all of them.

c) The above factorisations are valid for any density of the random variables a, b and c. Suppose now that the Bayesian network in Fig. 1 describes the joint distribution of a, b and c. Factorise the joint density $p(a, b, c)$ according to the Bayesian network. What does the Bayesian network tell us about $p(c|a, b)$? What is the interpretation of this in the example suggested above. Is it a reasonable assumption?

$P(a,b,c) =$
$P(c|b)P(b|a)P(b)$

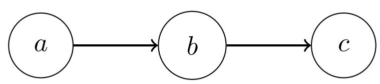

Figure 1: An illustration of the dependencies between a, b and c.

d) Repeat *b*) and *c*) for $p(x_1, x_2, \ldots, x_k)$ described by the Bayesian network in Fig. 2. What does the Bayesian network tell us about, e.g., $p(x_{k-1}|x_{k-2}, x_{k-3})$? What do you call a sequence with that property?

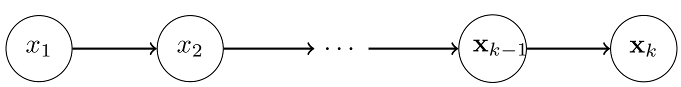

Figure 2: A Bayesian network representing the joint distribution $p(x_1, x_2, \ldots, x_k)$.

$$P(x_1, x_2 \cdots x_k) = P(x_k|x_{k-1})P(x_{k-1}|x_{k-2}) \cdots P(x_1)$$

$$= \left\{\prod_{i=2}^{k} P(x_i|x_{i-1})\right\} P(x_1)$$

Problem 2 – Sven goes downtown

In this problem we shall try to improve our understanding of the conditional independencies and dependencies in a state space model. To give it a more practical feeling, we will discuss a guy who goes shopping.

Problem setting: Sven is having lunch downtown and leaves a known restaurant at 12.55, which gives us prior knowledge about where he was at 13.00. We later obtain noisy GPS measurements of his position at 13.20 and 13.40; we denote these measurements as $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively. Suppose that we wish to figure out where Sven was at 13.00, 13.20 and 13.40 and that we denote his positions at those times as $\mathbf{x}_0$, $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively.

Model assumptions: Let us assume that his motion can be described as a random walk:

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \mathbf{q}_{k-1} \tag{1}$$

where $\mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\sigma_q^2)$ and that the GPS-measurements can be modelled as

$$\mathbf{y}_k = \mathbf{x}_k + \mathbf{r}_k \tag{2}$$

where $\mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\,\sigma_r^2)$, for $k = 1$ and $k = 2$. Let us also assume the conventional dependence structure for state space models, illustrated in Fig. 3.

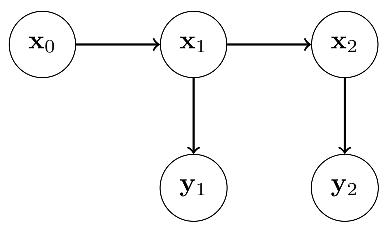

Figure 3: An illustration of the dependency structure in the filtering setting.

For the following densities, point out which of the variables that we condition on that can be left out and **motivate the simplifications in terms of Sven's movement and the GPS information**:

a) $p(\mathbf{x}_2|\mathbf{x}_1, \cancel{\mathbf{x}_0})$

b) $p(\mathbf{x}_2|\mathbf{x}_1, \cancel{\mathbf{y}_1})$

c) $p(\mathbf{y}_1|\mathbf{x}_1, \cancel{\mathbf{x}_0}, \cancel{\mathbf{x}_2})$ → 他要去哪

d) $p(\mathbf{x}_1|\mathbf{x}_0, \mathbf{y}_1)$ $\propto p(x_1, x_0, y_1) = p(y_1|x_1)\,p(x_1|x_0)\,p(x_0)$

e) $p(\mathbf{x}_1|\mathbf{x}_0, \mathbf{y}_1, \mathbf{x}_2, \cancel{\mathbf{y}_2})$

It is generally desirable to express densities of interest in terms of densities that we can model (in this case the motion and measurement models, e.g., $p(\mathbf{x}_1|\mathbf{x}_0)$ and $p(\mathbf{y}_1|\mathbf{x}_1)$).

f) Try to express the densities in *d*) and *e*) in terms of measurement and motion models. You can ignore normalisation factors, that is, factors that do not depend on $\mathbf{x}_1$.

Problem 3 – predicting Sven's movement using Chapman-Kolmogorov

In this problem we will try to understand the Chapman-Kolmogorov equation by applying it a couple of toy examples.

In the previous example, suppose that we instead saw Sven stepping out of a specific restaurant precisely at 13.00 such that our prior knowledge is

$$p(\mathbf{x}_0) = \delta(\mathbf{x}_0 - \hat{\mathbf{x}}_0) \tag{3}$$

where $\hat{\mathbf{x}}_0$ is the position of the restaurant.

a) What do you think that the predicted density $p(\mathbf{x}_1)$ is? An intuitive formulation using Sven's movement in the city is desirable.

   A fundamental question, that we discussed briefly during the lecture, is what happens to our uncertainties. Do they increase or decrease?

b) Use the Chapman-Kolmogorov equation to verify your guess. If your guess was wrong you should of course use the Chapman-Kolmogorov equation to correct it.

   For simplicity, you can assume that $\mathbf{x}_0$ and $\mathbf{x}_1$ are scalar variables and use for instance $\hat{x}_0 = 5$ (if you know that Sven always walks along the same street, a scalar is sufficient to describe his position).

   Hint: In case you do not remember the details of the Chapman-Kolmogorov equation, here it is:

$$p(\mathbf{x}_k \big| \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k \big| \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \big| \mathbf{y}_{1:k-1}) \, d\mathbf{x}_{k-1}. \tag{4}$$