Probabilistic Graphical Models in Computer Vision

Lecture: Dr. Tao Wu
Exercises: Yuesong Shen, Zhenzhang Ye
Summer Semester 2019

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 4

Room: 02.09.023

Wednesday, 05.06.2019, 12:15 - 14:00

Exact Inference (12+6 Points)

Exercise 1 (4 Points)**.** Firstly, draw one possible factor graph for each Markovf Random Fields shown as following. Then write the corresponding factorization and independence of following 4 Markov Random Fields:

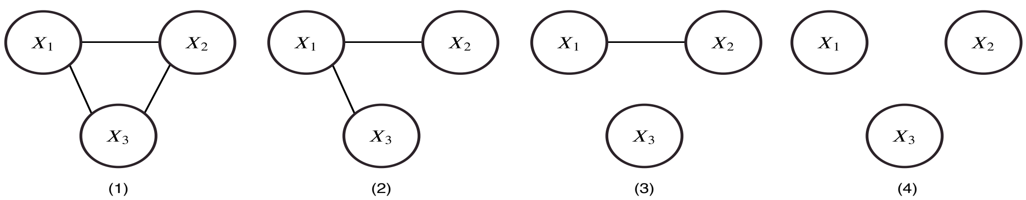


Figure 1: Exercise 1

Solution. Actually, each MRF has several factor graphs. Here we show one possible factor graph for each. The corresponding factor graphs are:

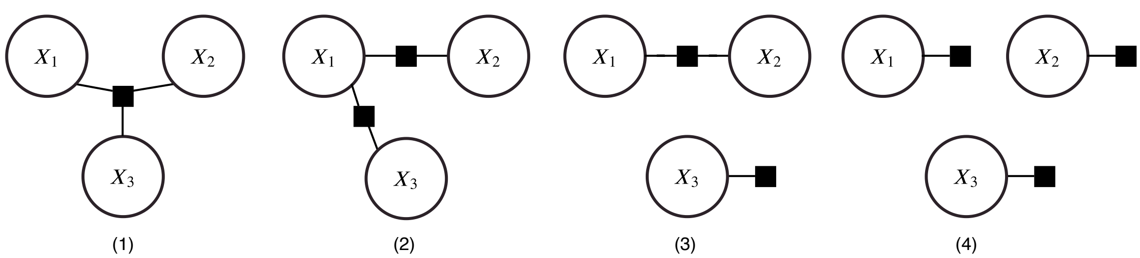


Figure 2: Exercise 2

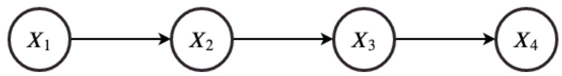
1. Factorization: $\frac{\phi(x_1,x_2,x_3)}{\sum_x (x_1,x_2,x_3)}$. Independence: None.

2. Factorization: $\frac{\phi(x_1,x_2)\phi(x_2,x_3)}{\sum_x \phi(x_1,x_2)\phi(x_2,x_3)}$. Independence: $X_2 \perp X_3 | X_1$.

3. Factorization: $\frac{\phi(x_1,x_2)\phi(x_3)}{\sum_x \phi(x_1,x_2)\phi(x_3)}$. Independence: $X_3 \perp (X_1, X_2)$.

4. Factorization: $\frac{\phi(x_1)\phi(x_2)\phi(x_3)}{\sum_x \phi(x_1)\phi(x_2)\phi(x_3)}$. Independence: pairwise independent.

Exercise 2 (4 Points)**.** Consider following directed chain of 4 random variables, each variable can take n number of values. Assume we want to evaluate the probability that X_4 takes on value x_4, *i.e.* $P(X_4 = x_4)$.



1. If we use straightforward probabilistic description, how many entries are required for a full joint probability table regarding n? How many operations are required ragarding n? (Use Big O notation).

2. Write down the factorization of $P(x_1, x_2, x_3, x_4)$ and explain how we can use it to simplify the computation.

3. How many operations do we need now regarding n? (Use Big O notation).

Solution. We first write down the formula for computing $P(X_4)$:

$$P(x_4) = \sum_{x_1,x_2,x_3} P(x_1, x_2, x_3, x_4) \tag{1}$$

1. If we write down the table directly, each variable has n number of values and we have 4 variables. The table will be n^4. To commpute the marginal distribution, we need to sum over other 3 variables, which involves $O(n^3)$ addition operations.

2. $P(x_1, x_2, x_3, x_4) = P(x_1)P(x_2|x_1)P(x_3|x_2)P(x_4|x_3)$. Now we plug it into above equation:

$$\begin{aligned}
P(x_4) &= \sum_{x_1,x_2,x_3} P(x_1)P(x_2|x_1)P(x_3|x_2)P(x_4|x_3) \\
&= \sum_{x_2,x_3} P(x_3|x_2)P(x_4|x_3) \sum_{x_1} P(x_1)P(x_2|x_1) \\
&= \sum_{x_2,x_3} P(x_2)P(x_3|x_2)P(x_4|x_3) \\
&= \sum_{x_3} P(x_3)P(x_4|x_3)
\end{aligned}$$

3. The operations are $O(n^2)$ now. Because for each variable, we need to marginalized it by considering the node and its child.

Exercise 3 (4 Points)**.** Using the same idea from previous exercise, consider following problem, assume each variable has n number of values:

1. Using variable elimination with the order x_1, x_2, x_4, x_3 to compute $P(x_5)$. How many operations do we need regarding n?

2. What if we eliminate with the order x_3, x_1, x_2, x_4 to compute $P(x_5)$? How many operations do we need reagarding n?

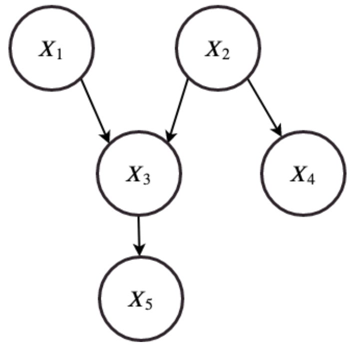


Figure 3: Exercise 3

Solution. To compute the marginal distribution of x_5 and consider the graph, we have :

$$p(x_5) = \sum_{x_1,x_2,x_3,x_4} p(x_1,x_2,x_3,x_4,x_5) = \sum_{x_1,x_2,x_3,x_4} p(x_1)p(x_2)p(x_3|x_1,x_2)p(x_4|x_2)p(x_5|x_3)$$

1. Follow the elimination order, we get:

$$\begin{aligned} p(x_5) &= \sum_{x_2,x_3,x_4} p(x_2)p(x_4|x_2)p(x_5|x_3)\sum_{x_1} p(x_1)p(x_3|x_1,x_2) \\ &\propto \sum_{x_2,x_3,x_4} \phi_1(x_2,x_3)p(x_2)p(x_4|x_2)p(x_5|x_3) \\ &= \sum_{x_3,x_4} p(x_5|x_3)\sum_{x_2} \phi_1(x_2,x_3)p(x_2)p(x_4|x_2) \\ &\propto \sum_{x_3,x_4} \phi_2(x_3,x_4)p(x_5|x_3) \\ &\propto \sum_{x_4} \phi_3(x_4,x_5) \end{aligned}$$

   In this case, each summation requires at most $O(n^3)$ operations.

2. If we start with x_3, we will get:

$$p(x_5) = \sum_{x_1,x_2,x_4} p(x_1)p(x_2)p(x_4|x_2)\sum_{x_3} p(x_3|x_1,x_2)p(x_5|x_3)$$

   which gives us $\phi(x_1,x_2,x_5)$ over 3 variables, which would require $O(n^4)$ operations. This might be more diffuclt when the graph becomes more complicated.

Exercise 4 (6 Points)**.** Here we use a simple factor graph to practice Belief Propagation. Consider following factor graph: assume designate node x_3 as the root, use the Belief Propagation to verify $\tilde{P}(x_2) \propto \sum_{x_1,x_3,x_4} \tilde{P}(x_1,x_2,x_3,x_4)$, where $\tilde{P}$ is the unnormalized probability.

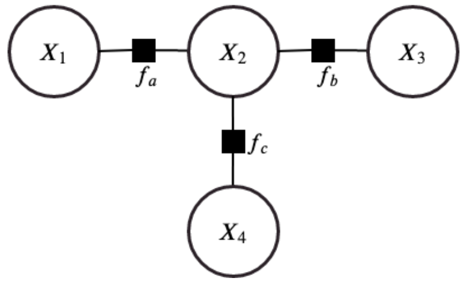


Figure 4: Exercise 4

Solution. Define $E_a(x_1, x_2) = -\log f_a(x_1, x_2)$ and so on. Starting from two leaf nodes x_1, x_4, we have following sequence of six messages:

$$
\begin{aligned}
q_{x_1 \to f_a}(x_1) &= 0 \\
r_{f_a \to x_2}(x_2) &= \log \sum_{x_1} \exp(-E_a(x_1, x_2)) \\
q_{x_4 \to f_c}(x_4) &= 0 \\
r_{f_c \to x_2}(x_2) &= \log \sum_{x_4} \exp(-E_c(x_2, x_4)) \\
q_{x_2 \to f_b}(x_2) &= r_{f_a \to x_2}(x_2) + r_{f_c \to x_2}(x_2) \\
r_{f_b \to x_3}(x_3) &= \log \sum_{x_2} \exp(-E_b(x_2, x_3) + q_{x_2 \to f_b}(x_2))
\end{aligned}
$$

Now, we propagate messages from the root node out to the leaf nodes:

$$
\begin{aligned}
q_{x_3 \to f_b}(x_3) &= 0 \\
r_{f_b \to x_2}(x_2) &= \log \sum_{x_3} \exp(-E_b(x_2, x_3)) \\
q_{x_2 \to f_a}(x_2) &= r_{f_b \to x_2}(x_2) + r_{f_c \to x_2}(x_2) \\
r_{f_a \to x_1}(x_1) &= \log \sum_{x_2} \exp(-E_a(x_1, x_2) + q_{x_2 \to f_a}(x_2)) \\
q_{x_2 \to f_c}(x_2) &= r_{f_a \to x_2}(x_2) + r_{f_b \to x_2}(x_2) \\
r_{f_c \to x_4}(x_4) &= \log \sum_{x_2} \exp(-E_c(x_2, x_4) + q_{x_2 \to f_c}(x_2))
\end{aligned}
$$

Then, we can evaluate the marginals:

$$
\begin{aligned}
\tilde{P}(x_2) &\propto \exp(r_{f_a \to x_2}(x_2) + r_{f_b \to x_2}(x_2) + r_{f_c \to x_2}(x_2)) \\
&\propto \sum_{x_1} \sum_{x_3} \sum_{x_4} f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4) \\
&\propto \sum_{x_1, x_3, x_4} \tilde{P}(x_1, x_2, x_3, x_4)
\end{aligned}
$$