Assignment 2

a) The assignments have to be done individually.

b) The assignments have to be answered in English.

c) The answers have to be uploaded to Toledo trough "Assignments", not using any other Toledo tool and not by email.

d) Clearly indicate your name and student number in the uploaded answer.

e) Indicate the time that you have (approximately) spend on the assignment. This will not be taken into account in any way for the quotation but will give us an idea of the load of the assignments. We estimate a load of maximally two hours and half per assignment.

f) You will be rewarded for correct answers, not for the format of your answer.

g) The name of the uploaded document should be **rXXXXXXX.pdf** where you replace XXXXXXX with your student number.

h) **Use the template of the solution sheet to submit your solution.** In case no printer is available, feel free to mimic this template by hand.

Markov Networks and the Ising Model

In this assignment you will study a problem that you also saw in class: recovering pixels. We will not study an entire image but focus on a sub-grid of 4×4 pixels (cf. Figure 1). The original image is represented by the pixels $\{X_1, \ldots, X_{16}\}$. Each of the pixels has two possible values $\{-1, 1\}$ and the image that we want to display are denoted by the pixels $\{Y_1, \ldots, Y_{16}\}$. In the first five questions we acquaint ourselves with the problem and move on to the reconstructing of a lost pixel in the sixth question.

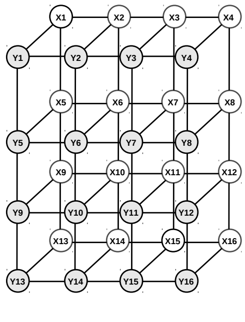

Figure 1: Markov network of pixels in image

Consider the following function over the variables $\mathcal{X} = \{X_1, \dots, X_{16}\}$ and $\mathcal{Y} = \{Y_1, \dots, Y_{16}\}$:

$$\begin{aligned} &E(X_1, \dots, X_{16}, Y_1, \dots, Y_{16}) \\ &= \alpha_1 \sum_{X_i \sim X_j} E_1(X_i, X_j) + \alpha_2 \sum_{Y_i \sim Y_j} E_2(Y_i, Y_j) + \alpha_3 \sum_i E_3(X_i, Y_i) \end{aligned} \tag{1}$$

In physics (where the Ising model originates) the function is also called energy function. Note how the energy function decomposes into the functions E_1, E_2, and E_3, which now only depend on two of the 32 variables. The links between the nodes in Figure 1 represent these binary couplings.
Energy functions relate in the following fashion to probability distributions:

$$P(\mathcal{X}, \mathcal{Y}) = P(X_1, \dots X_{16}, Y_1, \dots, Y_{16}) \propto e^{-\frac{E(\mathcal{X}, \mathcal{Y})}{T}} \tag{2}$$

1. Let us now consider the Markov network given in Figure 1 for which we write down the following probability distribution in terms of potentials ϕ_1, ϕ_2, and ϕ_3:

$$\begin{aligned} P(\mathcal{X}, \mathcal{Y}) &= P(X_1, \dots X_{16}, Y_1, \dots, Y_{16}) \\ &= \frac{1}{Z} \left(\prod_{X_i \sim X_j} \phi_1(X_i, X_j) \right) \left(\prod_{Y_i \sim Y_j} \phi_2(Y_i, Y_j) \right) \left(\prod_{i \in \{1, \dots, 16\}} \phi_3(X_i, Y_i) \right) \end{aligned} \tag{3}$$

   where Z is a normalization constant.

   Express ϕ_1, ϕ_2 and ϕ_3 in terms of E_1, E_2 and E_3.

2. Assume an Ising model as described on Slide 7 of the lecture notes. Express $E_1(X_i, X_j)$, $E_2(Y_i, Y_j)$ and $E_3(X_i, Y_j)$ in terms of X_i, X_j, Y_i and Y_j.

3. We now observe the following pixel values:

$$\begin{aligned} \{&X_1 = -1, X_2 = 1, X_3 = 1, X_4 = 1, \\ &X_5 = -1, X_6 = -1, X_7 = 1, X_8 = 1, \\ &X_9 = -1, X_{10} = 1, X_{11} = -1, X_{12} = 1, \\ &X_{13} = 1, X_{14} = -1, X_{15} = -1, X_{16} = 1\} \end{aligned} \tag{4}$$

$$\begin{aligned} \{&Y_1 = -1, Y_2 = 1, Y_3 = 1, Y_4 = 1, \\ &Y_5 = -1, Y_7 = 1, Y_8 = 1, \\ &Y_9 = -1, Y_{10} = 1, Y_{11} = -1, Y_{12} = 1, \\ &Y_{13} = 1, Y_{14} = -1, Y_{15} = -1, Y_{16} = 1\} \end{aligned} \tag{5}$$

   What is the probability of pixel Y_6 taking the value 1?

   In order to answer this question, assume $\alpha_1 = 0$, $\alpha_2 = 1$, and $\alpha_3 = 1$. Furthermore, use the functions E_1, E_2, and E_3 from the second question with $T = 1$. Make sure you take into account conditional independence properties of Markov networks to simplifiy computations.

4. Assume now that $\alpha_1 \neq 0$. Would your answer to the previous question change? Why?

5. Assume again $\alpha_1 = 0$ and also that you did not observe the pixel Y_7. In other words you do not observe Y_6 nor Y_7. Would your answer to the third question change? Why?

6. This time you have the following observations:

$$
\begin{aligned}
&\{X_2 = 1, X_3 = 1, X_4 = 1, && (6)\\
&\ X_5 = -1, X_6 = -1, X_7 = 1, X_8 = 1, \\
&\ X_9 = -1, X_{10} = 1, X_{11} = -1, X_{12} = 1, \\
&\ X_{13} = 1, X_{14} = -1, X_{15} = -1, X_{16} = 1\} \\
&\{Y_2 = 1, Y_3 = 1, Y_4 = 1, && (7)\\
&\ Y_5 = -1, Y_6 = -1, Y_7 = 1, Y_8 = 1, \\
&\ Y_9 = -1, Y_{10} = 1, Y_{11} = -1, Y_{12} = 1, \\
&\ Y_{13} = 1, Y_{14} = -1, Y_{15} = -1, Y_{16} = 1\}
\end{aligned}
$$

   This means that the pixel X_1 is defect and we do not have any signal for it. In the image we want to display, we have to infer a value for pixel Y_1 without knowing pixel X_1. What is the most likely value of Y_1 (1 or -1)?

   In order to answer this question, assume $\alpha_1 = 1$, $\alpha_2 = 1$, and $\alpha_3 = 1$. Furthermore, use again the potentials obtained with the answer to the second question with $T = 1$.

Name: Mohammad Sadil Khan

Student ID: r0873969

Solution

Fill in each answer in the respective box. Intermediate computations can be done on a separate page.

1.

$\phi_1(X_k, X_l) = \exp\left(-\frac{\alpha_1}{T} E_1(x_k, x_l)\right)$

$\phi_2(Y_k, Y_l) = \exp\left(-\frac{\alpha_2}{T} E_2(y_k, y_l)\right)$

$\phi_3(Y_k, X_l) = \exp\left(-\frac{\alpha_3}{T} E_3(y_k, x_l)\right)$

2.

$E_1(x_i, x_j) = \begin{cases} \frac{1}{2}(x_i - x_j)^2 & \text{if } X_i \sim X_j \\ 0 \end{cases}$

$E_2(Y_i, Y_j) = \begin{cases} \frac{1}{2}(y_i - y_j)^2 & \text{if } Y_i \sim Y_j \\ 0 & \text{oth} \end{cases}$

$E_3(X_i, Y_j) = \begin{cases} \frac{1}{2}(x_i - y_j)^2 & \text{if } i = j \\ 0 & \text{oth} \end{cases}$

$A \sim B$ implies there's an edge between A & B in the markov network

3. $P(Y_6 = 1 \mid X_1, X_2, \dots X_{16}, Y_1, \dots Y_5, Y_7, \dots Y_{16}) = 0.881$

4. If $\alpha_1 \neq 0$, the answer won't change.

Since $Y_6 \perp\!\!\!\perp X_i \mid X \setminus \{Y_6, X_i\}$ $\forall i = 1, 2 - 4, 6, \dots 16, i \neq 5$, no other variables, except x_5 from set X will be used during final calculation.

So, $\phi_1(x_i, x_j)$ will never be used.

5. Yes the $P(Y_6 = 1)$ will change since $Y_6 \not\perp\!\!\!\perp Y_7 \mid X \setminus \{Y_6, Y_7\}$, Y_6 & Y_7 are conditionally dependent given $X \setminus \{Y_6, Y_7\}$

(Reason stated below)

6. $P(Y_1 = 1 \mid Y_2 = 1, Y_5 = -1) = 0.5$

1.

$$P(x,y) = \frac{1}{Z}\left(\prod_{x_i \sim x_j} \varphi_1(x_i, x_j)\right)\left(\varphi_2(y_i, y_j)\right)\left(\prod_{i \in \{1,2,\dots k\}} \varphi_3(\cancel{x}_i, y_j)\right)$$

Again

$$P(X,Y) = \frac{1}{Z} \cdot e^{-\frac{1}{T}\left(\alpha_1 \sum_{x_i \sim x_j} E_1(x_i, x_j) + \alpha_2 \sum_{y_i \sim y_j} E_2(y_i, y_j) + \alpha_3 \sum_i E_3(x_i, y_i)\right)}$$

$$= \frac{1}{Z} e^{-\frac{\alpha_1}{T}\sum_{x_i \sim x_j} E_1(x_i, x_j) - \frac{\alpha_2}{T}\sum_{y_i \sim y_j} E_2(y_i, y_j) - \frac{\alpha_3}{T}\sum_i E_3(x_i, y_i)}$$

$$= \frac{1}{Z} \prod_{x_i \sim x_j} \exp\left(-\frac{\alpha_1}{T} E_1(x_i, x_j)\right) \prod_{y_i \sim y_j} \exp\left(-\frac{\alpha_2}{T} E_2(y_i, y_j)\right) \prod_i \exp\left(-\frac{\alpha_3}{T} E_3(x_i, y_j)\right)$$

$$\varphi_1(x_i, x_j) = -\frac{\alpha_1}{T} E_1(x_i, x_j)$$

$3 \therefore \alpha_1 = 0, \alpha_2 = 1, \alpha_3 = 1, T$

$$\therefore E = \sum_{y_i \sim y_j} E_1(y_i, y_j) + \sum_{x_i \sim y_j} E(x_i, y_j)$$

$$= \sum_{y_i \sim y_j} \frac{1}{2}(y_i - y_j)^2 + \sum_{x_i \sim y_j} \frac{1}{2}(x_i - y_j)^2$$

$$P(x_1, x_2, \ldots x_{16}, y_1, y_2, \ldots y_{16}) = \frac{1}{Z} \prod_{y_i \sim y_j} \phi_2(y_i, y_j) \prod_{i=1}^{16} \phi_3(x_i, y_i)$$

$$= \frac{1}{Z} \prod_{y_i \sim y_j} e^{-\frac{1}{2}(y_i - y_j)^2} \prod_{i=1}^{16} e^{-\frac{1}{2}(x_i - y_i)^2}$$

Now,

$$P(y_6 \mid x_1, x_2 \ldots x_{16}, y_1, \ldots y_5, y_7, \ldots y_{16})$$

$$= P(y_6 \mid x_6, y_2, y_5, y_7, y_{10}) \quad \left(\text{since } P(x \mid X \setminus \{x\}) = P(x / ne(x))\right)$$

$$= \frac{P(x_6, y_2, y_5, y_6, y_7, y_{10})}{P(x_6, y_2, y_5, y_7, y_{10})}$$

$$= \frac{P(x_6, y_2, y_5, y_6, y_7, y_{10})}{\sum_{y_6} P(x_6, y_2, y_5, y_6, y_7, y_{10})}$$

Given the values of the variables.

$$P(Y_6=1 \mid X_6=-1, Y_2=1, Y_5=-1, Y_7=1, Y_{10}=1)$$

$$= \frac{e^{-\frac{1}{2}\{(Y_6-X_6)^2+(Y_6-Y_5)^2+(Y_7-Y_6)^2+(Y_6-Y_2)^2+(Y_6-Y_{10})^2\}}}{\sum\limits_{Y_6\in\{-1,1\}} \exp\left(-\frac{1}{2}\{(Y_6-X_6)^2+(Y_6-Y_5)^2+(Y_7-Y_6)^2+(Y_6-Y_2)^2+(Y_6-Y_{10})^2\}\right)}$$

$$= \frac{\exp\left(-\frac{1}{2}\{4+4\}\right)}{\exp\left(-\frac{1}{2}\cdot 8\right)+\exp\left(-\frac{1}{2}(3\times 4)\right)}$$

$$= \frac{\exp(-4)}{\exp(-4)+\exp(-6)} = 0.881$$

5. Let $\alpha_1 = 0$

If Y_7 wasn't observed.

$$P(Y_6=1 \mid X_6=-1, Y_2=1, Y_5=-1, Y_{10}=1)$$

$$= \frac{\sum_{Y_7} P(X_6=-1, Y_6=1, Y_2=1, Y_5=-1, Y_7, Y_{10}=1)}{\sum_{Y_6, Y_7} P(X_6=-1, Y_6, Y_2=1, Y_5=-1, Y_7, Y_{10}=1)}$$

$$= \frac{e^{-4}+e^{-6}}{e^{-4}+e^{-6}+e^{-6}+e^{-4}} = \frac{1}{2} = 0.5$$

6. If X_1 isn't observed. then

$$P(Y_1=1 \mid X \setminus \{Y_1\}) = P(Y_1=1 \mid X_1, Y_2=1, Y_5=-1)$$

$$= \frac{\sum_{X_1} P(Y_1=1, X_1, Y_2=1, Y_5=-1)}{\sum_{X_1, Y_1} P(Y_1, X_1, Y_2=1, Y_5=-1)}$$

$$= \frac{e^{-2}+e^{-4}}{e^{-2}+e^{-4}+e^{-4}+e^{-2}}$$

$$= 0.5$$