CV201 HW2

Hadad Yakir ID: 313250276
Bukra Barak ID: 315453100

December 2019

1 Some General Questions About MRFs

1.1 Problem 1

Part (i)

X_1 \ X_3	0	1
0	(1-θ)^2	(1-θ)θ
1	(1-θ)θ	θ^2

Part (ii)

$$Y_1 = X_1 = \begin{cases} 1 & \theta \\ 0 & 1-\theta \end{cases}$$

Part (iii)

$$Y_2 = X_1 + X_2 = \begin{cases} 2 & \theta^2 \\ 1 & 2(1-\theta)\theta \\ 0 & (1-\theta)^2 \end{cases}$$

Part (iv)

$$Y_3 = X_1 + X_2 + X_3 = \begin{cases} 3 & \theta^3 \\ 2 & 3(1-\theta)\theta^2 \\ 1 & 3(1-\theta)^2\theta \\ 0 & (1-\theta)^3 \end{cases}$$

Part (v)

$$p(Y_1 = y_1 | X_1 = x_1) = \begin{cases} 1 & if \;\; y_1 = x_1 \\ 0 & else \end{cases}$$

Part (vi)

$$p(y_2|x_1,\ x_2) = \begin{cases} 1 & if \quad y_2 = x_1 + x_2 \\ 0 & else \end{cases}$$

Part (vii)

$$p(y_2|y_1,\ y_2) = 1$$

Part (viii)

$$p(y_3|x_1,\ x_2) = \begin{cases} \theta & if \quad y_3 = x_1 + x_2 + 1 \\ 1-\theta & if \quad y_3 = x_1 + x_2 \\ 0 & else \end{cases}$$

Part (ix)
y_1 dont give more information if we already have x_1 so its the same as $viii$

$$p(y_3|x_1,\ x_2,\ y_1) = p(y_3|x_1,\ x_2)$$

Part (x)
Same as ix y_2 dont give more information if we have x_1 and x_2

$$p(y_3|x_1,\ x_2,\ y_1) = p(y_3|x_1,\ x_2)$$

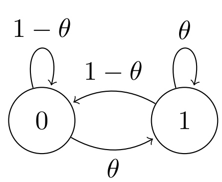
Part (xi)

$$p(y_3|x_1,\ x_2,\ x_3,\ y_2) = \begin{cases} 1 & if \quad y_3 = y_2 + x_3 \\ 0 & else \end{cases}$$

Part (xii)

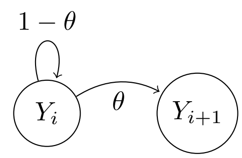
$$p(y_3|y_1,\ y_2,\ y_3) = 1$$

Part (xiii)
Yes, $p(x_i|x_{i-1}, x_{i-2}, \ldots, x_1) = p(x_i|x_{i-1}) = p(x_i)$ so by definition $(X_i)_{i=1}^n$ is markov chain.

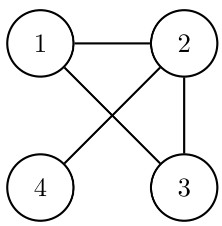


Part (xiv)

Yes, $p(y_i|y_{i-1}, y_{i-2}, \ldots, y_1) = p(y_i|y_{i-1})$ so again by definition $(Y_i)_{i=1}^n$ is markov chain. In general the state Y_i and the directed out edges look like that

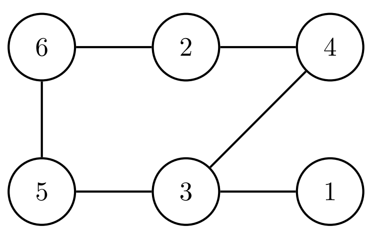


1.2 Problem 2



1.3 Problem 3

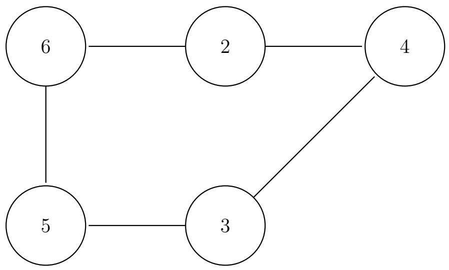
Part i)



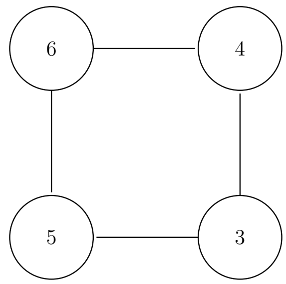
Part ii)

$$C = \{\{1,3\}, \{2,4\}, \{2,6\}, \{3,4\}, \{3,5\}, \{5,6\}\} \cup V_G$$

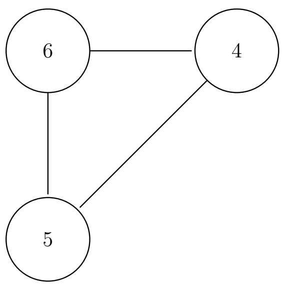
Part iii)



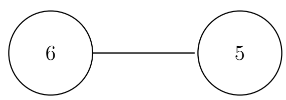
(a) G_{23456}



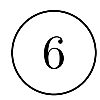
(b) G_{3456}



(c) G_{456}



(d) G_{56}



(e) G_6

Part iv)

(a) No, because there is still path from x_1 to x_6 without x_5.

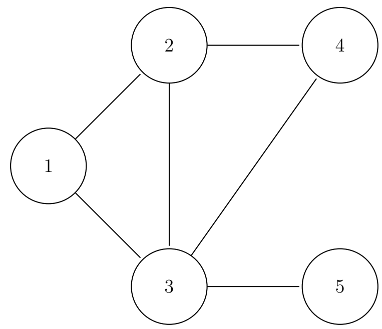
(b) No, x_1 and x_3 are neighbors so there are no subgroup of random variable of X that will separate it.

(c) Yes, x_4 and x_5 split x_1 and x_6 into to different subgroups.
Part v)
Property of procision matrix is that $Q_{ij} = 0$ iff $x_i \perp\!\!\!\perp x_j|_{x_1 x_2} X$

$$Q = \begin{bmatrix} - & 0 & - & 0 & 0 & 0 \\ 0 & - & 0 & - & 0 & - \\ - & 0 & - & - & - & 0 \\ 0 & - & - & - & 0 & 0 \\ 0 & 0 & - & 0 & - & - \\ 0 & - & 0 & 0 & - & - \end{bmatrix}$$

1.4 Problem 4

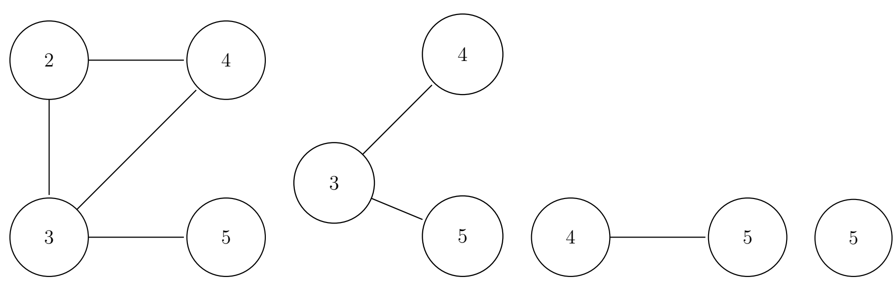
Part i)



Part ii)

$$C = \{\{1,2,3\},\{2,4\},\{3,4\},\{3,4\}\} \cup V_G$$

Part iii)



(a) G_{2345} (b) G_{345} (c) G_{45} (d) G_5

Part iv)

To see if x_1 and x_4 are conditionally independent given other variables, needs to check if this variables split x_1 and x_4 into other graphs.

a) No

b) Yes

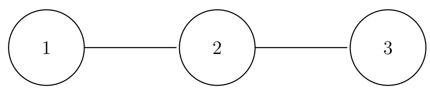
c) No

Part v)

$$Q = \begin{bmatrix} - & - & - & 0 & 0 \\ - & - & - & - & 0 \\ - & - & - & - & - \\ 0 & - & - & - & 0 \\ 0 & 0 & - & 0 & - \end{bmatrix}$$

1.5 5

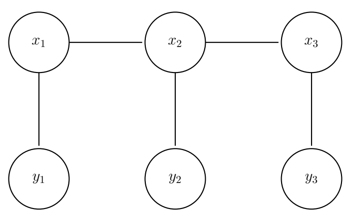
Part i)



$$C = \{\{1,2\},\{2,3\}\}$$

Part ii)

I'll draw G_{xy} to help me to solve it.



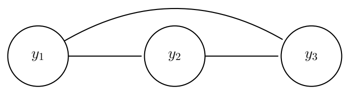
$$p(y_1, y_2, y_3|x_1, x_2, x_3) = \{y \textit{ conditionally independent given } x\} = \prod_{i=1}^{3} p(y_i|x_i) = \prod_{i=1}^{3} G(x_i, y_i)$$

Part iii)

$$p(x,y) = p(y|x)p(x) \propto \prod_{i=1}^{3} G(x_i, y_i) \prod_{c \in C} F_c(x_c)$$

Part iv)
I draw it in answer *ii*

Part v)



Part vi)

a) No

b) Yes, x_2 is enough

c) Yes

Part vii)

$$Q = \begin{bmatrix} - & - & 0 & - & 0 & 0 \\ - & - & - & 0 & - & 0 \\ 0 & - & - & 0 & 0 & - \\ - & 0 & 0 & - & 0 & 0 \\ 0 & - & 0 & 0 & - & 0 \\ 0 & 0 & - & 0 & 0 & - \end{bmatrix}$$

1.6 Problem 6

Part i)
The likelihood function is the same as above:

$$p(y|x) = \prod_{s \in S} p(y_s|x_s) = \prod_{s \in S} \frac{1}{\sqrt{(2\pi\sigma^2)}} e^{-\dfrac{(y_s - x_s)^2}{2\sigma^2}} \propto \prod_{s \in S} e^{-\dfrac{(y_s - x_s)^2}{2\sigma^2}}$$

Part ii)

$$
\begin{aligned}
\hat{x} &= \underset{x}{argmax}\ p(y|x) \\
&= \underset{x}{argmax}\ \prod_{s\in S} e^{-\dfrac{(y_s - x_s)^2}{2\sigma^2}} \\
&= \prod_{s\in S} \underset{x}{argmax}\ e^{-\dfrac{(y_s - x_s)^2}{2\sigma^2}} \\
&= \prod_{s\in S} \underset{x}{argmax}\ e^{-(y_s - x_s)^2} \\
&= \prod_{s\in S} {}_x argmin\ (y_s - x_s)^2
\end{aligned} \tag{1}
$$

$x_s \in \{-1, 1\}$ so to get the minimum of $y_s - x_s \Rightarrow x_s = sign(y_s)$

$$\hat{x} = sign(y)$$

Part iii)
It won't affect, because in our computation we didn't take into account the distribution of the prior but only its values. The way it can really affect its if the probability of one of the values that x_s can get will be 0, and then the estimator is not necessary because we always know what x is, which is just a boring case. For example if $p(x_s = 1) = 0$ then even if we get that $\hat{x}_i = 1$ we know it's the noise that caused it.

2 The Ising Model

2.1 Problem 7

$$E(x) = \sum_x z^{-1} x \exp(\beta \sum_{s\sim t} x_s x_t)$$

I want to define $\hat{x} = -x$ and $\epsilon(x)$ for the whole exponent for simpler writing.

$$\epsilon(x) = \exp(\beta \sum_{s\sim t} x_s x_t)$$

By opening $\epsilon(\hat{x})$ we see that:

$$\epsilon(\hat{x}) = \exp(\beta \sum_{s\sim t} \hat{x}_s \hat{x}_t) = \exp(\beta \sum_{s\sim t} -x_s \cdot -x_t) = \exp(\beta \sum_{s\sim t} x_s x_t) = \epsilon(x)$$

$$x\epsilon(x) + \hat{x}\epsilon(\hat{x}) = \epsilon(x)(x + \hat{x}) = \epsilon(x)(x - x) = 0$$

Because we sum over all x for every x its complement $\hat{x}$ exist to and they canceling each other which leaves us that $E(x) = 0$ $\square$

2.2 Problem 8

$$p\left(\begin{bmatrix}-1 & -1\\ 1 & 1\end{bmatrix}\right) = \frac{exp(\beta(1+1-1-1))}{z} = z^{-1}$$
$$p\left(\begin{bmatrix}-1 & -1\\ -1 & 1\end{bmatrix}\right) = \frac{exp(\beta(1-1+1-1))}{z} = z^{-1}$$

Same probabilities no matter the temp.

2.3 Problem 9

$$p\left(\begin{bmatrix}-1 & -1 & -1\\ 1 & 1 & 1\\ 1 & -1 & 1\end{bmatrix}\right) = \frac{exp(\beta(1+1+1+1-1-1-1+1-1-1-1+1))}{z} = z^{-1}$$
$$p\left(\begin{bmatrix}-1 & 1 & -1\\ -1 & 1 & -1\\ -1 & 1 & -1\end{bmatrix}\right) = \frac{exp(\beta(-1-1-1-1-1-1+1+1+1+1+1+1))}{z} = z^{-1}$$

Same probabilities no matter the temp.

2.4 Problem 10

If temp getting smaller it means the β grows, the bigger β than the probability to bigger stains is growing, because the sum of the neighbors give a bigger exponent and we know that exponent is a monotonic function.
The smaller β gets than it flat more and more the exponent which mean if temp grows than the distribution getting more and more closer to uniform distribution.

2.5 Problem 11

Let $\phi(x) = x$, that would be enough to simply give a better chance to 1's in x over -1's

2.6 Problem 12

As I mentioned in problem 10 the bigger temp the flatten the exponent function so in a situation of $temp \to \infty$ $p(x) \xrightarrow{temp\to\infty} z^{-1}e^{0}$ which is total uniform distribution and $X \sim U\{2^{64}\}$

3 Computer Exercise

3.1 Problem 14

We see that the covariance between different pixels has a downward trend as we expected from the theory part. Plus we see that as the further the pixel is the less those pixels are correlated.

```
    --- Computer Exercise 3 ---
Z_1 of 2x2 lattice = 121.23293134406595
Z_1.5 of 2x2 lattice = 40.922799092745386
Z_2 of 2x2 lattice = 27.048782764334526
    --- Computer Exercise 3 ---
```

Figure 3: Computer Exercise 3

```
    --- Computer Exercise 4 ---
Z_1 of 3x3 lattice = 365645.74913577037
Z_1.5 of 3x3 lattice = 10565.421983514265
Z_2 of 3x3 lattice = 2674.518123060087
    --- Computer Exercise 4 ---
```

Figure 4: Computer Exercise 4

```
    --- Computer Exercise 5 ---
Z_1 of 2x2 lattice = 121.23293134406595
Z_1.5 of 2x2 lattice = 40.922799092745386
Z_2 of 2x2 lattice = 27.048782764334526
    --- Computer Exercise 5 ---
```

Figure 5: Computer Exercise 5

```
    --- Computer Exercise 6 ---
Z_1 of 3x3 lattice = 365645.7491357704
Z_1.5 of 3x3 lattice = 10565.421983514265
Z_2 of 3x3 lattice = 2674.518123060087
    --- Computer Exercise 6 ---
```

Figure 6: Computer Exercise 6

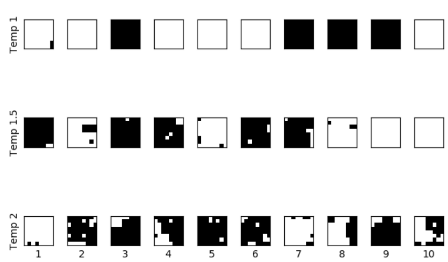


Figure 7: Computer Exercise 7

```
    --- Computer Exercise 8 ---
For temp = 1 => E(X11 x X22) = 0.9462
For temp = 1 => E(X11 x X88) = 0.8998
For temp = 1.5 => E(X11 x X22) = 0.9006
For temp = 1.5 => E(X11 x X88) = 0.8196
For temp = 2 => E(X11 x X22) = 0.8548
For temp = 2 => E(X11 x X88) = 0.7278
    --- Computer Exercise 8 ---
```

Figure 8: Computer Exercise 8

```
Total Running time 0:00:48.577340
```

Figure 9: Total Running time