ML-2

Homework Assigment-7

Shantanu Chandra
(12048461)

Q1.

1)

$$A_{new} = \underset{A}{\operatorname{argmax}}\; Q(\theta, \theta^{old})$$

or, $$\frac{\partial Q(\theta,\theta^{old})}{\partial A} = -\frac{1}{2}\sum_{n=2}^{N} \mathbb{E}_{z|\theta^{old}} \frac{\partial}{\partial A}\left[\frac{1}{2}(z_n - Az_{n-1})^T \Gamma^{-1}(z_n - Az_{n-1})\right]$$

$$= -\frac{1}{2}\sum_{n=2}^{N} \mathbb{E}_{z|\theta^{old}}\left[-2\Gamma^{-1}(z_n - Az_{n-1})\, z_{n-1}^T\right]$$

(using 88 of matrix cookbook : $\frac{\partial}{\partial A}(x-As)^T W (x-As) = -2W(x-As)s^T$)

or, $$\cancel{\Gamma^{-1}}\sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}[z_n z_{n-1}^T] = \cancel{\Gamma^{-1}} A \sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}[z_{n-1}z_{n-1}^T]$$

$$\boxed{A = \frac{\sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}[z_n z_{n-1}^T]}{\sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}[z_{n-1}z_{n-1}^T]}}$$

Ans.

$$\Gamma_{new} = \underset{\Gamma}{\operatorname{argmax}}\; Q(\theta,\theta^{old})$$

or, $$\frac{\partial Q(\theta,\theta^{old})}{\partial \Gamma} = -\frac{N-1}{2}\underbrace{\frac{\cancel{|\Gamma|}\,\Gamma^{-T}}{\cancel{|\Gamma|}}}_{\text{matrix cookbook 49}} - \mathbb{E}_{z|\theta^{old}}\frac{\partial}{\partial\Gamma}\left[\frac{1}{2}(z_n - Az_{n-1})^T\Gamma^{-1}(z_n - Az_{n-1})\right]$$

(using $\frac{\partial |X|}{\partial X} = |X|(X^{-1})^T$)

$$= -\frac{N-1}{2}\Gamma^{-1} - \mathbb{E}_{z|\theta^{old}}\left[-\frac{1}{2}\Gamma^{-1}(z_n - Az_{n-1})(z_n - Az_{n-1})^T\Gamma^{-1}\right]$$

(using matrix cookbook 61)

or, $$(N-1)\Gamma^{-1} = \Gamma^{-1}\,\mathbb{E}_{z|\theta^{old}}\left[(z_n - Az_{n-1})(z_n - Az_{n-1})^T\right]\Gamma^{-1}$$

Multiplying LHS & RHS by Γ from the right & the left:

$$\Gamma(N-1) = \mathbb{E}_{z|\theta^{old}}\left[(z_n - Az_{n-1})(z_n - Az_{n-1})^T\right]$$

$$\therefore \boxed{\Gamma = \frac{1}{N-1}\mathbb{E}_{z|\theta^{old}}\left[(z_n - Az_{n-1})(z_n - Az_{n-1})^T\right]}$$

Ans.

2) The steps for C^{new} are analogous to A^{new} & Σ is to Γ

$$C = \frac{\sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}[x_n z_n]}{\sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}[z_n z_n^T]}$$

Ans.

$$\Sigma = \frac{1}{N}\sum_{n=2}^{N}\mathbb{E}_{z|\theta^{old}}\left[(x_n - Cz_n)(x_n - Cz_n)^T\right]$$

Ans.

Q2.

a) 1. $X \quad Y$ 2. $X \rightarrow Y$ 3. $X \leftarrow Y$

[NOTE: $X \rightleftarrows Y$ is not a valid option as it is cyclic & Bayesian networks are acyclic]

b)
1. $p(X,Y) = p(X)\,p(Y)$
2. $p(X,Y) = p(X)\,p(Y|X)$
3. $p(X,Y) = p(Y)\,p(X|Y)$

c)
1. $p(Y|X) = \dfrac{p(X,Y)}{p(X)} = \dfrac{p(X)\cdot p(Y)}{p(X)} = p(Y)$ // Ans.
2. $p(Y|X) = \dfrac{p(X,Y)}{p(X)} = \dfrac{p(X)\,p(Y|X)}{p(X)} = p(Y|X)$ // Ans.
3. $p(Y|X) = \dfrac{p(X,Y)}{p(X)} = \dfrac{p(Y)\,P(X|Y)}{P(X)}$ // Ans.

d) perfect intervention do(X) causes all the incoming edges to X to be dropped as it is fixed now & hence nothing can cause it. In such a case:

1. $X \quad Y$ $p(Y|do(X)) = \dfrac{p(X,Y)}{P(X)} = \dfrac{p(X)\,p(Y)}{p(X)} = P(Y)$ // Ans.
2. $X \rightarrow Y$ $p(Y|do(X)) = \dfrac{p(X,Y)}{P(X)} = \dfrac{p(Y|X)\,p(X)}{P(X)} = p(Y|X)$ // Ans.
3. $X \quad Y$ $p(Y|do(X)) = \dfrac{p(X,Y)}{p(X)} = \dfrac{P(X)\,p(Y)}{P(X)} = p(Y)$ // Ans.

e) $p(Y|X)$ or $p(\text{cancer}\,|\,\text{smokes})$ is the case where we have the observation of people having cancer who smoked. Here there is a possibility of latent causes affecting this outcome (Eg: confounders), ie, factors that led to the person to smoke (by his will), might also be affecting the variable 'cancer'.

$p(Y|do(X))$ or $p(\text{cancer}\,|\,do(\text{smokes}))$ is the case where we have control over the smoking variable as we force this action on the subjects. In doing so, we remove all the other latent causes that were leading to smoking and might have also affected the outcome 'cancer' indirectly.

Q3.

1) $p(\text{recovery} \mid \text{drug}) = \frac{20}{40} = 0.5 = 50\%$

$p(\text{recovery} \mid \text{control}) = \frac{16}{40} = 0.4 = 40\%$

Based on just this information, we would recommend ANY new patient to take the drug as it has better recovery rate.

Ans.

2) $p(\text{recovery} \mid \text{drug, male}) = 60\%$

$p(\text{recovery} \mid \text{control, male}) = 70\%$

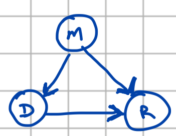
$p(\text{recovery} \mid \text{drug, female}) = 20\%$

$p(\text{recovery} \mid \text{drug, female}) = 30\%$

With the added information about the gender of the patient, we see that recovery rates are higher for the control group.
Thus, advice not to take drug for both male & female.

3) Based on new inform., I would advice not to take the drug irrespective of the gender. This does contradict the previous advice. This is because the recovery rate of men is much higher (irrespective of drug/control) than women in general. Since no. of males taking the drug is disproportionately higher, on average it biases the results even more towards taking drug. Though, this result should be inconclusive due to bias due to imbalance.

Q4.

a) 

Here, we look for nodes admissable for adjustment.
$S = \{M\}$ satisfies this as:

1) D, R ∉ S
2) no element of S is decendent of D.
3) There is one back-door path between D & R (incoming to D) that is blocked by S (M is a non-collider & is in S).

Thus, M is admissable for adjustment:

$$p(R \mid do(D)) = \sum_{M} p(R \mid M, D)\, p(M)$$

Ans.

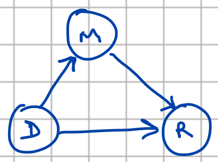
b) As shown above, empty set is not admissable for adjustment here, hence $p(R|do(D)) \neq p(R|D)$

c) As seen in Q3. 2), when the gender is taken into consideration <u>'not taking drug' is better</u>. We can see this as follows too:

$$p(R|do(D)) = \sum_M p(R|M,D)\, p(M)$$

$$= \frac{1}{2} \times \frac{18}{30} + \frac{1}{2} \times \frac{2}{10} = \underline{\underline{40\%}}$$

Q5.

a) 

Here, we look for nodes admissable for adjustment.
$S = \{\phi\}$ satisfies this as:
1) $D, R \notin S$
2) no element of S is decendent of D.
3) There are <u>no back-door</u> path between D & R (incoming to D).
Thus, <u>empty set is admissable</u> for adjustment.

$\therefore\ S = \{\phi\}$

$\therefore\ p(R|do(D)) = p(R|D)$

Ans.

b) As seen above, since the empty set is permissable, <u>yes this holds</u>.

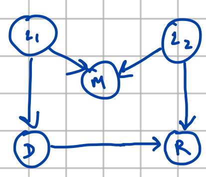
c) As seen in Q3. 1), it is <u>advisable to take drug</u> in this case (60% recovery rate).

Q6.

a) <u>L_1</u> can be the <u>economic background</u> of the individual. This affects D, ie, wealthier individuals are more likely to be able to afford drugs for treatment.

<u>L_2</u> can be the <u>genetics of the individual</u>. This can affect R, ie, some individuals with certain genetic traits may recover better.

<u>M</u> can be <u>no. of gold medals won</u>. L_1 & L_2 directly affect this (wealthier individuals with better genetics are more likely to win more). Also, M does not affect/cause anything else in the graph.

b) 

Here, we look for nodes admissable for adjustment.
$S = \{\phi\}$ satisfies this as:
1) $D, R \notin S$
2) no element of S is decendent of D.
3) There is <u>one back-door</u> path between D & R (incoming to D) that is blocked by S (M is a collidor in path & is not in S).

$\therefore\ S = \{\phi\}$

$\therefore\ p(R|do(D)) = p(R|D)$

Ans.

c) As seen above, since the empty set is permissable, yes this holds.

d) As seen in Q3. 1), it is advisable to take drug in this case (60% recovery rate).