HW 7 Name Changxin Miu
ID 11853018

P1

Find A^{new} and Γ^{new} that optimise

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

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

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

$$= \mathbb{E}_{z|\theta^{old}}\left[\Gamma^{-1}z_n z_{n-1} - \frac{1}{2}\Gamma^{-1}Az_n z_{n-1}^T + \Gamma^{-1}Az_{n-1}z_{n-1}^T\right] = 0$$

$$A^{new} = \left(\sum_{n=2}^{N}\mathbb{E}[z_n z_{n-1}^T]\right)\left(\sum_{n=2}^{N}\mathbb{E}[z_{n-1}z_{n-1}^T]\right)^{-1}$$

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

$$= -\frac{N-1}{2}(\Gamma^{-1}) - \mathbb{E}_{z|\theta^{old}}\left[-\frac{1}{2}\sum_{n=2}^{N}\Gamma^{-1}(z_n - Az_{n-1})(z_n - Az_{n-1})^T\Gamma^{-1}\right] = 0$$

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

$$\Gamma^{new} = \frac{1}{N-1}\sum_{n=2}^{N}\{\mathbb{E}[z_n z_n^T] - A^{new}\mathbb{E}[z_{n-1}z_n^T] - \mathbb{E}[z_n z_{n-1}^T]A^{new} + A^{new}\mathbb{E}[z_{n-1}z_{n-1}^T](A^{new})^T\}$$

For C^{new} and Σ^{new}

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

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

$$= \mathbb{E}_{z|\theta^{old}}\left[\Sigma^{-1}x_n z_n - \frac{1}{2}[\Sigma^{-1}Cz_n z_n^T + \Sigma^{-1}Cz_n^T z_n\right]$$

$$C^{new} = \left(\sum_{n=1}^{N}x_n\mathbb{E}[z_n^T]\right)\left(\sum_{n=1}^{N}\mathbb{E}[z_n z_n^T]\right)^{-1}$$

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

$$= -\frac{N}{2}(\Sigma^{-1}) - \mathbb{E}_{z|\theta^{old}}\left[-\frac{1}{2}\sum_{n=1}^{N}(-\Sigma^{-1}(x_n - Cz_n)(x_n - Cz_n)^T\Sigma^{-1}\right] = 0$$

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

$$\Sigma^{new} = \frac{1}{N}\sum_{n=1}^{N}\{x_n x_n^T - C^{new}\mathbb{E}[z_n]x_n^T - x_n\mathbb{E}[z_n^T]C^{new} + C^{new}\mathbb{E}[z_n z_n^T]C^{new}\}$$

Problem 2

(a) $X \quad Y$ (1) new effect $\qquad X \rightarrow Y$ (2) X causes Y $\qquad X \leftarrow Y$ (3) Y causes X

(b) a(1) $p(X,Y) = p(X)p(Y)$

a(2) $p(X,Y) = P(Y|X) \cdot p(X)$

a(3) $p(X,Y) = p(X|Y) \cdot p(Y)$

(c) a(1) $p(Y|X) = p(X)p(Y)$

a(2) $p(Y|X) = p(Y|X)$

a(3) $p(Y|X) = \dfrac{p(X|Y) \cdot p(Y)}{p(X)} = \dfrac{p(X|Y) \cdot p(Y)}{\sum_Y p(X,Y)} = \dfrac{p(X|Y) \cdot p(Y)}{\sum_Y p(X|Y)p(Y)}$

(d) a(1) $p(Y|do(X)) = p(Y)$

a(2) $p(Y|do(X)) = p(Y|X)$

a(3) $p(Y|do(X)) = p(Y)$

(e) $p(Y|X)$ means given the person is smoking what is his probability of getting the lung cancer.

$p(Y|do(X))$ means the probability for a person to get lung cancer after he/she starts to smoke, there is an intervention action in this probability.

P3

1a. Drug: $\frac{\text{Recovery}}{\text{Total}} = \frac{20}{40} = 50\%$

no drug: $\frac{\text{Recovery}}{\text{Total}} = \frac{16}{40} = 40\%$

1b: Yes, the recovery rate is higher for the treatment group

2a M: Drug: $\frac{18}{30} = 60\%$

no drug: $\frac{7}{10} = 70\%$

F: Drug: $\frac{2}{10} = 20\%$

no drug: $\frac{9}{30} = 30\%$

2b: I would not advise patients to take drugs since the recovery rate is higher if the patient does not take the drug.

3 As the recovery rate for female and male patients declines after taking the drug I would advice a diseased patient with unknown gender to not taking the drug. This is a contradiction to my ealier advice.

4. a $S = \{M\}$. In this case, M is not descendant of $\{D, R\}$. Also M blocks all back path from R to D.

$P(R \mid do(D)) = \int P(R \mid D, m)\, p(m)\, dm = \sum_M P(R \mid D, M)\, P(M)$ M is admissible for adjustment for the causal effect of R to D.

4b. $P(R \mid D) = \sum_M P(R, M \mid D) = \sum_M P(R \mid D, M) \cdot P(D \mid M)$

$P(D \mid M) \neq P(M)$ Hence $P(R \mid do(D)) \neq P(R \mid D)$

4c. $P(R \mid do(D)) = P(R \mid D=1, M=m) \cdot P(M=m) + P(R \mid D=1, M=f) \cdot P(M=f)$

$= 60\% \times 50\% + 20\% \times 50\% = 40\%$

$P(R \mid do(\neg D)) = P(R \mid D=0, M=m) \cdot P(M=m) + P(R \mid D=0, M=f) \cdot P(M=f)$

$= 70\% \times 50\% + 30\% \times 50\% = 50\%$

~~Acc~~ According to the above calculation, the probability of recovery is lower when patients are taking drugs. Hence, I would not advise to take drugs.

~~4d.~~

5a. $S = \{\emptyset\}$ ~~$S=\{M\}$~~ $\emptyset$ could ~~not~~ be admissible for adjustment from R to D.

$P(R \mid d(D)) =$ ~~$P(R \mid D, M) = \sum_M P(R$~~

$\sum_M P(R \mid D, M)\, P(M \mid D) = P(R \mid D)$

5b. Yes. $P(R \mid do(D)) = P(R \mid D)$

5c. ~~I would~~

$P(R \mid do(D)) = P(R \mid D=1, M=m)\, P(M=m \mid D=1) + P(R \mid D=1, M=f) \cdot P(M=f \mid D=1)$

$= 0.6 \times 0.75 + 0.2 \times 0.25 = 0.5$

$P(R \mid do(\neg D)) = P(R \mid D=0, M=m)\, P(M=m \mid D=0) + P(R \mid D=0, M=f) \cdot P(M=f \mid D=0)$

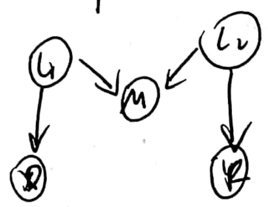
$= 70\% \times 0.25 + 30\% \times 0.75 = 0.4$

$P(R \mid do(D)) > P(R \mid do(\neg D))$

In this case, I would advise patients with unknown gender to take drugs.

6a. R, D, M represents recovery, taking drugs and gender. L_1 can be the mood of the patient which infuences his ability to absord the drug and it also influences his physical condition (gender). L_2 could be the genetic element to influcence the gender of recovery ability of the patient.

6c. Applying the second rule of do-calculus, we could observe that the graph change to



, which is obvionse that $(D \perp\!\!\!\perp R)_{G_{\underline{D}}}$ and given M is the collider for this path

Therefore $P(R \mid do(D)) = P(R \mid D)$

~~the causal effect for R, D according to back door criterion~~

~~L_1, L_2 are admissible for adjustmen for R, D~~

6b: ~~$P(R \mid do(D)) = \sum_{L_1} P(R \mid D, L_2) P(L_2)$~~

~~$P(R \mid do(D)) = \sum_{L_2} P(R \mid D, L_1) P(L_1)$~~

~~$P(R \mid D) =$~~

$P(R \mid do(D)) = P(R \mid D)$ According to the second rule of do-calculus. Since $(D \perp\!\!\!\perp R)_{G_{\underline{D}}}$.

6d: $P(R \mid do(D)) = 0.5$

$P(R \mid do(\neg D)) = 0.4$

Therefore, the drug should be recommend to an unknown gender patient

Problem 4

(1) $P(W) = \cancel{p(W|R\ \ p(W,R,S) = \sum_R p(W|R) + p(}\ \sum_R \sum_S p(W, U_R=1, U_S=1)$

$\cancel{= \sum_R p(W|R)\,p(R{=}1) + \sum_S p(W|S)\cdot p(S{=}1)}$

$= \cancel{p(W|R=1)\times 0.7 + p(W|U_S=1)\times 0.4}$

$= p(W, U_R=1, U_S=1) + p(W|U_R=1)\,p(U_R) + p(W|U_S=1)\,p(U_S{=}1)$ $p(U_S=0)$ $p(U_R=0)$

$= p(W|U_R=1, U_S=1)\,p(U_R=1, U_S=1) + p(W|U_R=1)\,p(U_R=1)\,p(U_S=0) + p(W|U_S=1)\,p(U_S=1)\,p(U_R=0)$

$= 1\times 0.7\times 0.4 + 1\times 0.7\times(1-0.4) + 1\times(1-0.7)\times 0.4$

$= 0.82$

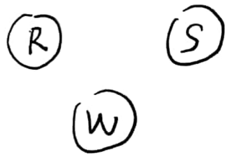
(2) $P(R=1|W=1) = P(R=1, S=1|W=1) + P(R=1, S=0|W=1)$

$$= \frac{P(W=1|R=1,S=1)\cdot P(R=1,S=1)}{P(W=1)} + \frac{P(W=1|R=1,S=0)\cdot P(R=1,S=0)}{P(W=1)}$$

$$= \frac{0.4\times 0.7}{0.82} + \frac{\cancel{0.7\times}\ 0.7\times 0.6}{0.82} = \frac{0.7}{0.82} = 0.8537$$

(3) The causal effect can not be identified, there might be a latent $\underbrace{\text{variable}}_{}$ confounds such as humidity has positive effect on both R and W.

(4)



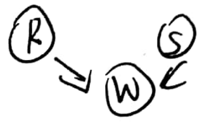
(5) ~~P(R, S | do~~ According to the above graph, we can get that

$P(R|do(W=w)) = P(R)$

$\begin{cases} P(R=0) = 0.3 \\ P(R=1) = 0.7 \end{cases}$

$P(R|W) = \begin{cases} P(R=1|W=0) = 0 \\ P(R=0|W=0) = 1 \\ P(R=1|W=1) = 0.85 \\ P(R=0|W=1) = 0.15 \end{cases}$

(6)



(7) $S = \{\emptyset\}$ for the ~~eas~~ causal relationship between S and.

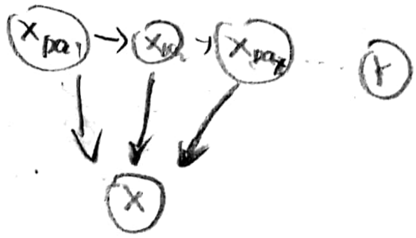
Hence $P(W|do(S=s)) = P(W|S=s)$

They are different from $P(W)$.

5c

Problem 5

$\cdots h, M=m \mid D=1) + p(R \mid D=1, M=f) \cdot p(M=f \mid D=1)$



1. $$P(Y \mid do(X), x_{pa(x)}) = \frac{P(Y, X_{pa(x)} \mid do(X))}{P(X_{pa(x)} \mid do(X))}$$
$$= \frac{P(Y) \prod_{x_i \in X_{pa(x)}} P(X_i)}{\prod_{x_i \in X_{pa(x)}} P(X_i)}$$
$$= P(Y)$$
$$= \frac{P(Y)\, P(X \mid X_{pa(x)})\, P(X_{pa(x)})}{P(X \mid X_{pa(x)})\, P(X_{pa(x)})}$$
$$= \frac{P(Y, X, X_{pa(x)})}{P(X, X_{pa(x)})}$$
$$= P(Y \mid X, X_{pa(x)})$$

2. $$P(Y \mid do(X)) = \int P(Y \mid do(X), X_{pa(X)})\, P(X_{pa(x)} \mid do(X))\, d X_{pa(x)}$$
$$= \int P(Y \mid X, X_{pa(x)}) \prod_{x_i \in X_{pa(x)}} P(X_i)\, d X_{pa(X)}$$
$$= \int P(Y \mid X, X_{pa(x)})\, P(X_{pa(x)})\, d X_{pa(X)}$$