ML2 Home Assignment 7

id: 12179078

Volodymyr Medentsiy

PROBLEM 2

a) I

$$X \longrightarrow Y$$

II

$$X \longleftarrow Y$$

III

$$X \qquad Y$$

b) I $p(X,Y) = p(X) \cdot p(Y|X)$

II $p(X,Y) = p(Y)\, p(X|Y)$

III $p(X,Y) = p(X)\, p(Y)$

c) I $p(Y|X) = \dfrac{p(X,Y)}{p(X)}$

II $p(Y|X) = \dfrac{p(X,Y)}{p(X)} = \dfrac{p(X|Y)\, p(Y)}{p(X)}$

III $p(Y/X) = \dfrac{p(X,Y)}{p(X)} = \dfrac{p(X)\, p(Y)}{p(X)} = p(Y)$

d) I $p(Y|do(X)) = \left[X \text{ has no parents}\right] = p(Y|X)$

II $p(Y|do(X)) = p(Y)$, Y has no parents.

III $p(Y|do(X)) = p(Y)$, Y has no parents.

e) Y - lung cancer , X - smoking.

$p(Y|X)$ - is a probability that a person will have cancer if he smokes. $p(Y|do(X))$ – corresponds to probability that a person will have cancer if we intentionally force him to smoke.

The key difference is that when we calculate $p(Y|do(X))$ we disregard all other variables which could have caused smoking of a person and also be the reason of cancer.

PROBLEM 3

1a. Recovery rate / Drug = 50%

Recovery rate / No drug = 40%

1.b. I would advise taking the drug

2a.

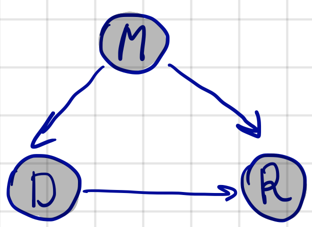
Males	Recovery	No recovery	Total	Recovery rate
Drug	18	12	30	60..%
No drug	7	3	10	70.%
Total	25	15	40	

Females	Recovery	No recovery	Total	Recovery rate
Drug	2	8	10	20..%
No drug	9	21	30	30..%
Total	11	29	40	

2b. I do not know, these are very confusing results. Probably I would advise against taking the drug for both male and females.

3. If the gender is unknown I would advise taking the drug.

4.



$p(R|do(D))$ – ? To block the path $R \to M \to D$ we need to include

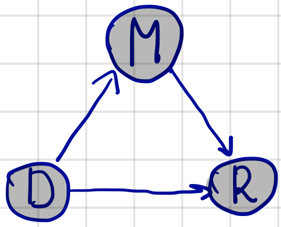
Min S, thus we obtain:

$$p(R|do(D)) = \sum_m p(R|D,m)\,p(m)$$

4b. no $p(R|do(D)) \neq p(R|D)$

4c. I would advise not to take the drug, because recovery rate considering gender is lower among those who took the drug.

5.



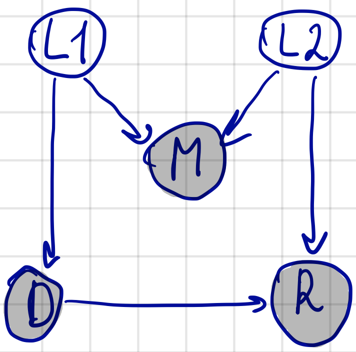
a. $S = \{\emptyset\}$, because D has no parents. Thus $p(R|do(D)) =$

$$p(R|D) = \sum_m p(R|D,m)\,p(m|D)$$

b. $p(R|do(D)) = p(R|D)$

c. Then I would advise to take the drug, as the probability in general population to recover if taking the drug is higher.

6.



Let M denote the strenght of immune system of a person.

L2 denotes how rich a person is. L1 – where the person lives, assume in some areas there is a limited supply of a drug and instead of using original drug people are prescribed with its cheaper version.

a. M is collider, thus blocks path from R to D with incoming edge from $L1$ to D $\Rightarrow S = \{\emptyset\}$

$$P(R|do(D)) = P(R|D) = \iiint p(R|D, m, l_2, l_1)\, p(m|l_2, l_1)\, p(l_1)\, p(l_2)\, dm\, dl_2\, dl_1$$

b. yes $p(R|do(D)) = p(R|D)$

c. As in the previous example, I would advise to take the drug.

PROBLEM 1

$$p(z_n|z_{n-1}) = \mathcal{N}(z_n|Az_{n-1}, \Gamma)$$

$$p(x_n|z_n) = \mathcal{N}(x_n|Cz_n, \Sigma)$$

$$p(z_1) = \mathcal{N}(z_1|\mu_0, V_0)$$

$$\ln p(X, Z|\theta) = \ln p(z_1|\mu_0, V_0) + \sum_{n=1}^{N} \ln p(z_n|z_{n-1}, A, \Gamma) + \sum_{n=1}^{N} \ln p(X_n|z_n, C, \Sigma)$$

E-step: find $Q(\theta, \theta^{old}) = E_{Z|Z^{old}}(\ln p(X, Z|\theta))$

1. Find A^{new}, Γ^{new}, s.t. optimize:

$$Q(\theta, \theta^{old}) = -\frac{N-1}{2} \ln|\Gamma| - E_{Z|\theta^{old}}\left(\frac{1}{2}\sum_{n=2}^{N}(z_n - Az_{n-1})^T \Gamma^{-1}(z_n - Az_{n-1})\right) + \text{const.}$$

$$\frac{\partial Q(\theta, \theta^{old})}{\partial A} = E_{Z|\theta^{old}} \sum_{n=2}^{N} \Gamma^{-1}(z_n - Az_{n-1})\, z_{n-1}^T = 0 \Rightarrow$$

$$\Rightarrow E_{Z|\theta^{old}}\left(\sum_{n=2}^{N} \Gamma^{-1} z_n z_{n-1}^T - \sum_{n=2}^{N} \Gamma^{-1} A z_{n-1} z_{n-1}^T\right) = 0 \Rightarrow$$

$$\Rightarrow E_{Z|\theta^{old}} \sum_{n=2}^{N} \Gamma^{-1} A z_{n-1} z_{n-1}^T = E_{Z|\theta^{old}} \sum_{n=2}^{N} \Gamma^{-1} z_n z_{n-1}^T \quad \Big| \times \Gamma, \text{ right-side multiplication.}$$

$$\Rightarrow E_{Z|\theta^{old}} \sum_{n=2}^{N} A z_{n-1} z_{n-1}^T = E_{Z|\theta^{old}} \sum_{n=2}^{N} z_n z_{n-1}^T \quad \Big| \times \left(\sum_{n=2}^{N} z_{n-1} z_{n-1}^T\right)^{-1}, \text{ left-side}$$

$$\Rightarrow A^{new} = E_{z|\theta^{old}}\left(\sum_{n=2}^{N} z_n z_{n-1}^T\right) \cdot E_{z|\theta^{old}}\left(\sum_{n=2}^{N} z_{n-1} z_{n-1}^T\right)$$

$$\frac{\partial Q(\theta,\theta^{old})}{\partial \Gamma} = \left[\begin{array}{l} \dfrac{\partial \ln|X|}{\partial X} = (X^T)^{-1} \\ \dfrac{\partial a^T X^{-1} b}{\partial X} = -X^{-T} a b^T X^{-T} \end{array}\right] = -\frac{N-1}{2} \cdot \Gamma^{-1} -$$

$$- E_{z|\theta^{old}}\left(\frac{1}{2}\sum_{n=2}^{N} -\Gamma^{-1}(z_n - A z_{n-1})(z_n - A z_{n-1})^T \Gamma^{-1}\right) = 0 \Rightarrow$$

$$\Rightarrow \frac{N-1}{2} \cdot \Gamma^{-1} = E_{z|\theta^{old}}\left(\frac{1}{2}\sum_{n=2}^{N} \Gamma^{-1}(z_n - A z_{n-1})(z_n - A z_{n-1})^T \Gamma^{-1}\right) \Rightarrow$$

$$\Rightarrow \Gamma = \frac{1}{N-1}\sum_{n=2}^{N} E_{z|\theta^{old}}(z_n - A z_{n-1})(z_n - A z_{n-1})^T$$

2. $Q(\theta,\theta^{old}) = -\frac{N}{2}\ln|\Sigma| - E_{z|\theta^{old}}\left(\frac{1}{2}\sum_{n=1}^{N}(x_n - C z_n)^T \Sigma^{-1}(x_n - C z_n)\right)$

$$\frac{\partial Q(\theta,\theta^{old})}{\partial C} = E_{z|\theta^{old}}\sum_{n=1}^{N}\Sigma^{-1}(x_n - C z_n) z_n^T = 0$$

using result in previous step:

$$C^{new} = E_{z|\theta^{old}}\left(\sum_{n=1}^{N} x_n z_n^T\right) \cdot E_{z|\theta^{old}}\left(\sum_{n=1}^{N} x_n z_n^T\right)$$

$$\frac{\partial Q(\theta,\theta^{old})}{\partial \Sigma} = -\frac{N}{2}\Sigma^{-1} - E_{z|\theta^{old}}\left(\frac{1}{2}\sum_{n=1}^{N} -\Sigma^{-1}(x_n - C z_n)(x_n - C z_n)^T \Sigma^{-1}\right) = 0$$

$$\Rightarrow \Sigma = \frac{1}{N}\sum_{n=1}^{N} E_{z|\theta^{old}}(x_n - C z_n)(x_n - C z_n)^T$$