Homework 6

JD Gallego Posada, *University of Amsterdam* 22/05/2017

Problem 1

Collaborator: D Kianfar

$$1)\ \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) = -\frac{N-1}{2}\log|\boldsymbol{\Gamma}| - \frac{1}{2}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\sum_{n=2}^{N}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^\mathsf{T}\boldsymbol{\Gamma}^{-1}(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})\right] + K$$

$$= -\frac{N-1}{2}\log|\boldsymbol{\Gamma}| - \frac{1}{2}\sum_{n=2}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_n^\mathsf{T}\boldsymbol{\Gamma}^{-1}\mathbf{z}_n - 2\mathbf{z}_n^\mathsf{T}\boldsymbol{\Gamma}^{-1}\mathbf{A}\mathbf{z}_{n-1} + \mathbf{z}_{n-1}^\mathsf{T}\mathbf{A}^\mathsf{T}\boldsymbol{\Gamma}^{-1}\mathbf{A}\mathbf{z}_{n-1})\right] + K$$

$$\frac{\partial\mathcal{Q}}{\partial\mathbf{A}} = -\frac{1}{2}\sum_{n=2}^{N} -2\boldsymbol{\Gamma}^{-1}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_n^\mathsf{T}\mathbf{z}_{n-1}\right] + 2\boldsymbol{\Gamma}^{-1}\mathbf{A}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_{n-1}^\mathsf{T}\mathbf{z}_{n-1}\right] = 0$$

$$\mathbf{A}^{\text{new}} = \left(\sum_{n=2}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_n^\mathsf{T}\mathbf{z}_{n-1}\right]\right)\left(\sum_{n=2}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_{n-1}^\mathsf{T}\mathbf{z}_{n-1}\right]\right)^{-1}$$

$$\frac{\partial\mathcal{Q}}{\partial\boldsymbol{\Gamma}} = -\frac{N-1}{2}\boldsymbol{\Gamma}^{-1} - \frac{1}{2}\left(-\boldsymbol{\Gamma}^{-1}\sum_{n=2}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^\mathsf{T}\right]\boldsymbol{\Gamma}^{-1}\right) = 0$$

$$\boldsymbol{\Gamma}^{\text{new}} = \frac{1}{N-1}\sum_{n=2}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})(\mathbf{z}_n - \mathbf{A}\mathbf{z}_{n-1})^\mathsf{T}\right]$$

$$2)\ \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) = -\frac{N}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\sum_{n=2}^{N}(\mathbf{x}_n - \mathbf{C}\mathbf{z}_n)^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\mathbf{x}_n - \mathbf{C}\mathbf{z}_n)\right] + K$$

$$= -\frac{N}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{n=1}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{x}_n^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{x}_n - 2\mathbf{x}_n^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{C}\mathbf{z}_n^\mathsf{T} + \mathbf{z}_n^\mathsf{T}\mathbf{C}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\mathbf{C}\mathbf{z}_n)\right] + K$$

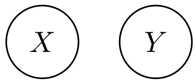
$$\frac{\partial\mathcal{Q}}{\partial\mathbf{C}} = -\frac{1}{2}\sum_{n=2}^{N} -2\boldsymbol{\Sigma}^{-1}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{x}_n^\mathsf{T}\mathbf{z}_n\right] + 2\boldsymbol{\Sigma}^{-1}\mathbf{C}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_n^\mathsf{T}\mathbf{z}_n\right] = 0$$

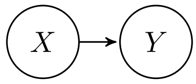
$$\mathbf{C}^{\text{new}} = \left(\sum_{n=1}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{x}_n^\mathsf{T}\mathbf{z}_n\right]\right)\left(\sum_{n=1}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[\mathbf{z}_n^\mathsf{T}\mathbf{z}_n\right]\right)^{-1}$$

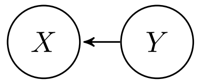
$$\frac{\partial\mathcal{Q}}{\partial\boldsymbol{\Sigma}} = -\frac{N}{2}\boldsymbol{\Sigma}^{-1} - \frac{1}{2}\left(-\boldsymbol{\Sigma}^{-1}\sum_{n=1}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[(\mathbf{x}_n - \mathbf{C}\mathbf{z}_n)(\mathbf{x}_n - \mathbf{C}\mathbf{z}_n)^\mathsf{T}\right]\boldsymbol{\Sigma}^{-1}\right) = 0$$

$$\boldsymbol{\Sigma}^{\text{new}} = \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}_{\mathbf{z}|\boldsymbol{\theta}^{\text{old}}}\left[(\mathbf{x}_n - \mathbf{C}\mathbf{z}_n)(\mathbf{x}_n - \mathbf{C}\mathbf{z}_n)^\mathsf{T}\right]$$

Problem 2

a) (i) 

(ii) 

(iii) 

b) (i) $p(X,Y) = p(X)p(Y)$

(ii) $p(X,Y) = p(X)p(Y|X)$

(ii) $p(X,Y) = p(X|Y)p(Y)$

c) (i) $p(Y|X) = p(Y)$

(ii) $p(Y|X) = p(Y|X)$

(iii) $p(Y|X) = \frac{p(X|Y)p(Y)}{\int p(Y)p(X|Y)dY}$

d) Note that $(Y \perp\!\!\!\perp X)_{\mathcal{G}_{\underline{X}}}$ holds in (i) and (ii). Therefore, by the *action/observation exchange* rule:

(i) $p(Y|\text{do}(X)) = p(Y|X) = P(Y)$

(ii) $p(Y|\text{do}(X)) = p(Y|X)$

Besides, $(Y \perp\!\!\!\perp X)_{\mathcal{G}_{\overline{X}}}$ in (iii). Therefore, by the *ignoring actions* rule:

(iii) $p(Y|\text{do}(X)) = p(Y)$

e) $p(\text{cancer}|\text{smokes})$ represents the probability that someone has cancer given that we observe that the person smokes. On the other hand, $p(\text{cancer}|\text{do}(\text{smokes}))$ represents the probability that someone has cancer given that we force that person to smoke.
Conditioning implies an update in our beliefs by incorporating observational evidence, while *do* represents what would happen if we intervened the studied system.

Problem 3

1a) $p(R|D=1) = \frac{20}{40} = 50\%$ $\quad$ $p(R|D=0) = \frac{16}{40} = 40\%$

1b) Disregarding the difference between causation and correlation, yes.

2a) Males: $p(R|D=1) = \frac{18}{30} = 60\%$ $\quad$ $p(R|D=0) = \frac{7}{10} = 70\%$

Females: $p(R|D=1) = \frac{2}{10} = 20\%$ $\quad$ $p(R|D=0) = \frac{9}{30} = 30\%$

2b) No in both cases.

3) Before, without knowing the gender, the recommendation was to take the drug. Now, independent of the gender, the patient should not take it.

4a) Since $S = \{M\}$ blocks the only back-door path, it is admissible for adjustment, and therefore $p(R|\text{do}(D)) = \int p(R|D,M)p(M)dM$.

4b) $p(R|D) = \int p(R|D,M)p(M|D)dM$, which is in general different from $p(R|\text{do}(D))$.

4c) Assuming model (i), no, since:

$$\begin{aligned} p(R|\text{do}(D=1)) &= p(R|D=1,M=1)p(M=1) + p(R|D=1,M=0)p(M=0) \\ &< p(R|D=0,M=1)p(M=1) + p(R|D=0,M=0)p(M=0) \\ &= p(R|\text{do}(D=0)) \end{aligned}$$

5a) Since there are no back-door paths, $\emptyset$ is admissible for adjustment, and $p(R|\text{do}(D)) = p(R|D)$.

5b) Yes.

5c) Assuming model (ii), yes.

6a) L_1: insulin production, M: blood pressure, L_2: glucagon production.

6b) $\emptyset$ blocks the only back-door path. Thus, $p(R|\text{do}(D)) = p(R|D)$.

6c) Yes.

6d) Assuming model (iii), yes.

Problem 4

1) We have $p(R=1)=0.7$, $p(S=1)=0.4$ and $p(W|R,S)=\chi_{W=R\vee S}$. Besides the joint distribution factorizes as $p(R,S,W)=p(R)p(S)p(W|R,S)$.

2) $$p(R=1|W=1)=\frac{p(R=1,W=1)}{p(W=1)}=\frac{\sum_S p(R=1,S,W=1)}{\sum_{R,S} p(R,S,W=1)}=\frac{0.42+0.28}{0.12+0.42+0.28}=\frac{0.7}{0.82}=0.853$$

3) No. The observational evidence updates our beliefs about R via conditioning.

4) $R=E_R$, $S=E_S$, $W=1$. Besides, we have $E_R \perp\!\!\!\perp E_S$.

5) $p(R|\text{do}(W=w))=p(R)$. This is in general different from $p(R|W=w)$.

6) $R=E_R$, $S=s$, $W=R\vee S$. Besides, we have $E_R \perp\!\!\!\perp E_S$.

7) $(W \perp\!\!\!\perp S)_{\mathcal{G}_{\underline{S}}}$ holds. Therefore, $p(W|\text{do}(S=s))=p(W|S=s)$, in general different from $p(W)$.

Problem 5

1) $$p(Y|\text{do}(X),\mathbf{X}_{pa(X)})=\frac{p(Y,\mathbf{X}_{pa(X)}|\text{do}(X))}{p(\mathbf{X}_{pa(X)}|\text{do}(X))}=\frac{\frac{P(Y,X,\mathbf{X}_{pa(X)})}{p(X|\mathbf{X}_{pa(X)})}}{\frac{P(X,\mathbf{X}_{pa(X)})}{p(X|\mathbf{X}_{pa(X)})}}=\frac{P(Y,X,\mathbf{X}_{pa(X)})}{P(X,\mathbf{X}_{pa(X)})}=p(Y|X,\mathbf{X}_{pa(X)})$$

2) Note that $p(\mathbf{X}_{pa(X)}|\text{do}(X))=p(\mathbf{X}_{pa(X)})$.

$$p(Y|\text{do}(X))=\int p(Y|\text{do}(X),\mathbf{X}_{pa(X)})p(\mathbf{X}_{pa(X)}|\text{do}(X))d\mathbf{X}_{pa(X)}=\int p(Y|X,\mathbf{X}_{pa(X)})p(\mathbf{X}_{pa(X)})d\mathbf{X}_{pa(X)}$$