به نام خدا

دانشگاه تهران

دانشکده مهندسی برق و کامپیوتر

درس مدل های مولد عمیق

تمرین اول

کیانا هوشانفر

۸۱۰۱۰۱۳۶۱

Problem ①:

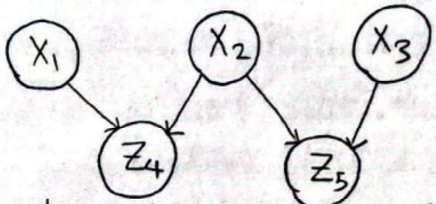
1. $P(X_2, X_3 | Z_5) = ?$

$$P(X_2 X_3 | Z_5) = \frac{P(X_2 X_3, Z_5)}{P(Z_5)} \overset{Z_5 = X_2 \oplus X_3}{=\!=\!=} \frac{P(X_2 X_3)}{P(Z_5)}$$

$$\Rightarrow \begin{cases} P(Z_5=0) = P(X_2=0, X_3=0) + P(X_2=1, X_3=1) = (1-q)^2 + q^2 \\ P(Z_5=1) = P(X_2=0, X_3=1) + P(X_2=1, X_3=0) = 2q(1-q) \end{cases}$$

$$\xRightarrow{\boxed{Z_5=0}} \begin{cases} P(X_2=0, X_3=0 | Z_5=0) = \frac{(1-q)^2}{(1-q)^2+q^2} \\ P(X_2=1, X_3=1 | Z_5=0) = \frac{q^2}{(1-q)^2+q^2} \\ P(X_2=1, X_3=0 | Z_5=0) = 0 \\ P(X_2=0, X_3=1 | Z_5=0) = 0 \end{cases}$$

$$\xRightarrow{\boxed{Z_5=1}} \begin{cases} P(X_2=0, X_3=0 | Z_5=1) = 0 \\ P(X_2=1, X_3=1 | Z_5=1) = 0 \\ P(X_2=1, X_3=0 | Z_5=1) = \frac{1}{2} \\ P(X_2=0, X_3=1 | Z_5=1) = \frac{1}{2} \end{cases}$$

2. 

$$\xrightarrow{X_1, X_2, X_3} \begin{cases} X_1 \to X_1=0 \to 1-q \\ \quad \hookrightarrow X_1=1 \to q \\ X_2 \to X_2=0 \to 1-q \\ \quad \hookrightarrow X_2=1 \to q \\ X_3 \to X_3=0 \to 1-q \\ \quad \hookrightarrow X_3=1 \to q \end{cases}$$

$Z_4 \to$

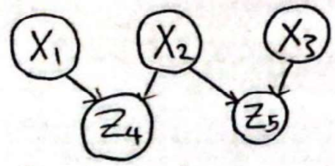
X_1	X_2	$Z_4=0$	$Z_4=1$
0	0	1	0
0	1	0	1
1	0	0	1
1	1	1	0

احتمال‌ها برای Z_4 مانند Z_5 بدست می‌آید ولی بجای X_1، X_3 داریم.

$Z_5 \to$ احتمال‌ها برای Z_5 در بالا محاسبه شده است

X_2	X_3	$Z_5=0$	$Z_5=1$
0	0	1	0
0	1	0	1
1	0	0	0
1	1	1	0

$$\Rightarrow \begin{array}{c|c} X_i=0 & X_i=1 \\ \hline 1-q & q \end{array}$$



Independence relations:

$$\begin{cases} X_1 \perp (X_2, X_3, Z_5) \\ X_2 \perp (X_1, X_3) \\ X_3 \perp (X_1, X_2, Z_4) \end{cases}$$

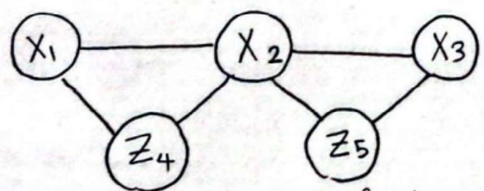
$$\begin{cases} X_1 \perp (X_2, X_3) \mid Z_5 \\ X_1 \perp (X_2, Z_5) \mid X_3 \\ X_3 \perp (X_1, X_2) \mid Z_4 \\ X_3 \perp (X_2, Z_4) \mid X_1 \end{cases}$$

$$\begin{cases} X_1 \perp X_2 \mid (X_3, Z_5) \\ X_1 \perp Z_5 \mid (X_2, X_3, Z_4) \\ X_2 \perp X_3 \mid (X_1, Z_4) \\ X_3 \perp Z_4 \mid (X_1, X_2, Z_5) \\ Z_4 \perp Z_5 \mid (X_1, X_2, X_3) \end{cases}$$

$$\begin{cases} X_1 \perp (X_3, Z_5) \mid (X_2, Z_4) \\ X_3 \perp (X_1, Z_4) \mid (X_2, Z_5) \\ Z_4 \perp (X_3, Z_5) \mid (X_1, X_2) \\ Z_5 \perp (X_1, Z_4) \mid (X_2, X_3) \end{cases}$$

$$(X_1, Z_4) \perp (X_3, Z_5) \mid X_2$$

3.



: undirected graphical model

Potential functions:

maximal
max - cliques:

$$\varphi_{142}(X_1, Z_4, X_2) = P(X_1)P(X_2)P(Z_4 \mid X_1 X_2) \overset{OR}{=} P(X_1)P(Z_4 \mid X_1 X_2)$$

$$\varphi_{253}(X_2, Z_5, X_3) = P(X_3)P(Z_5 \mid X_2 X_3) \overset{OR}{=} P(X_3)P(X_2)P(Z_5 \mid X_2 X_3)$$

Independence relations:

$$\begin{cases} X_1 \perp (X_3, Z_5) \mid (Z_4, X_2) \\ Z_4 \perp (X_3, Z_5) \mid (X_1, X_2) \end{cases}$$

$$\begin{cases} (X_1, Z_4) \perp X_3 \mid (X_2, Z_5) \\ (X_1, Z_4) \perp Z_5 \mid (X_2, X_3) \end{cases}$$

$$(X_1, Z_4) \perp (X_3, Z_5) \mid X_2$$

$$\begin{cases} X_1 \perp X_3 \mid (X_2, Z_4, Z_5) \\ X_1 \perp Z_5 \mid (X_2, X_3, Z_4) \\ Z_4 \perp Z_5 \mid (X_1, X_2, X_3) \\ Z_4 \perp X_3 \mid (X_1, X_2, Z_5) \end{cases}$$

4.

(I) $q \in \{0, 1\}$:

اگر $q \in \{0,1\}$ باشد، X_1, X_2, X_3 متغیرهای تصادفی ثابت می‌شوند که در نتیجه استقلال‌های $Z_4 \perp X_1$ و $Z_5 \perp X_3$ را نتیجه می‌دهد.
(" " " " " Z_5, Z_4)

(II) $0 < q < 1$:

for independence: $P(Z_5=1) = P(Z_5=1 \mid X_3=0) = P(Z_5=1 \mid X_3=1)$

$$P(Z_5=1) = \underbrace{P(Z_5=1 \mid X_3=0)}_{q} \underbrace{P(X_3=0)}_{1-q} + \underbrace{P(Z_5=1 \mid X_3=1)}_{1-q} \underbrace{P(X_3=1)}_{q}$$

$$\Rightarrow P(Z_5=1) = 2q(1-q)$$

استقلال‌های مورد نظر به ازای $0<q<1$ زمانی برقرار است که $q = \frac{1}{2}$ باشد.
marginal independencies از روی گراف‌های قسمت‌های قبل قابل مشاهده نیست.

Problem (2):

→ Figure 1:

1. season ⊥ chills → False → مستقل نیست چون FlU، give نشده و مسیر block نشده
2. Season ⊥ chills | Flu → True → مسیر بین دو تا متغیر توسط FlU، block شده.
3. Season ⊥ Headache | Flu → False → این دو متغیر نمی‌توانند لزوما مستقل باشند چون مسیر Season → Dehy → Headach وجود دارد که چون Dehy، given نشده،
4. Season ⊥ Headache | Flu, Dehydration → True → FlU و Dehy، given شدند و مسیرها توسط آن‌ها block شدند
5. Season ⊥ Nausea | Dehydration → False → بین این دو متغیر مسیر دیگری وجود دارد که این متغیرها given نشدند. Season → FlU → headach → Dizziness → Neasea
6. Season ⊥ Nausea | Dehydration, Headache → True → مسیرهای موجود توسط headach, Dehy block شده‌اند در FlU → head ← Dehy / v structure داریم.
7. FlU ⊥ Dehydration → False → (Flu ← Season → Dehy) → season، given نشده.
8. Flu ⊥ Dehydration | Season, Headache → False → V-structure به دلیل given شدن headach به وجود آمده
9. Flu ⊥ Dehydration | Season → True → مسیر توسط season block شده، همچنین به دلیل given نشدن head، V-structure نداریم
10. FlU ⊥ Dehydration | Season, Nausea → False → به دلیل given شدن nausea، V-structure داریم و چون dehy به Dizziness ربط پیدا می‌کند و headache give نشده تا مسیر block شود.
11. Chills ⊥ Nausea → False → متغیری given نشده تا مسیر block شود.
12. Chills ⊥ Nausea | Headache → False → در Dehy → Head ← FlU، به دلیل given شدن head V-structure داریم

→ Figure 2:

1. $P(S,F,D,C,H,N,Z) = P(S)P(F|S)P(D|S)P(C|F)P(H|F,D)P(Z|H)P(N|D,Z)$

2. $= \frac{1}{Z}\varphi_1(S)\varphi_2(F)\varphi_3(D)\varphi_4(C)\varphi_5(H)\varphi_6(N)\varphi_7(Z)\varphi_8(S,F)\varphi_9(S,D)$
$\varphi_{10}(F,C)\varphi_{11}(F,H)\varphi_{12}(D,H)\varphi_{13}(D,N)\varphi_{14}(H,Z)\varphi_{15}(N,Z)$

→ Table 1:

1. $P(Flu = True) = \sum_{season} P(Flu = True, Season = \{winter, Summer\})$

$= \sum_{s} P(Flu = True | S = s) P(S = s)$

$= P(Flu = True | S = winter) P(S = winter) + P(Flu = True | S = summer) P(S = summer)$

= ۰/۴ × ۰/۵ + ۰/۱ × ۰/۵ = $\boxed{۰/۲۵}$

2. $P(Flu = True | Season = Winter) = \boxed{۰/۴}$

3. $P(Flu = True | Season = winter, Headache = True) = \frac{P(F = True, S = w, H = true)}{P(S = W, H = True)}$

$= \frac{\sum_d P(F = True, S = w, H = True, D = d) \rightarrow P(H=T|F=T, D=d)P(F=T|S=W)P(D=d|S=W) \times P(S=W)}{\sum_{f,d} P(F = f, S = W, H = True, D = d) \rightarrow P(H=T|F=f, D=d)P(F=f|S=W)P(D=d|S=W)P(S=W)}$

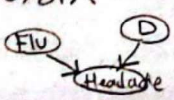
$= \frac{\overbrace{۰/۹×۰/۴×۰/۱×۰/۵}^{۰/۰۱۸} + \overbrace{۰/۸×۰/۴×۰/۹×۰/۵}^{۰/۱۴۴}}{\underbrace{۰/۹×۰/۴×۰/۱×۰/۵}_{۰/۰۱۸} + \underbrace{۰/۸×۰/۴×۰/۹×۰/۵}_{۰/۱۴۴} + \underbrace{۰/۸×۰/۶×۰/۱×۰/۵}_{۰/۰۲۴} + \underbrace{۰/۳×۰/۶×۰/۹×۰/۵}_{۰/۰۸۱}}$

$= \boxed{۰/۶۱}$

4. $P(Flu = True | Season = winter, Headache = True, Dehydration = True)$.

$= \frac{P(F = true, S = W, H = True, D = True) \rightarrow P(H=T|F=T, D=T)P(F=T|S=W)P(D=T|S=W)P(S=W)}{\sum_f P(F = f, S = W, H = T, D = T) \rightarrow \sum_f P(H=T|F=f, D=T)P(F=f|S=W)P(D=T|S=W)P(S=W)}$

$= \frac{۰/۹×۰/۴×۰/۱×۰/۵ = ۰/۰۱۸}{\underbrace{۰/۹×۰/۴×۰/۱×۰/۵}_{۰/۰۱۸} + \underbrace{۰/۸×۰/۶×۰/۱×۰/۵}_{۰/۰۲۴}} = \boxed{۰/۴۳}$

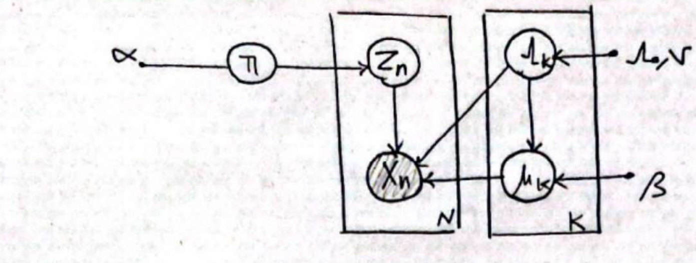
5. 

dehydration باعث کاهش احتمال رخ دادن Flu میشود. چون makes sense
D باعث H میشود.

Scanned with CamScanner

Problem ③:

E روی hidden variable ها می‌گیرم داخل E ظاهر شدن



$\Rightarrow \pi, Z_1, \dots, Z_n$
$X_1, \dots, X_n$
$\Lambda_1, \dots, \Lambda_K$
$\mu_1, \dots, \mu_K$

$$\log P(\Lambda_j \mid \text{other}) = E_{-j}[\log P(\text{all})] + \text{const}$$

$$\log P(\text{all}) = \underbrace{\log P(\pi)}_{\text{const}} + \underbrace{\sum_{i=1}^{N} \log P(Z_i \mid \pi)}_{\text{const}} + \underbrace{\sum_{i=1}^{N} \log P(X_i \mid Z_i, \Lambda_{Z_i}, \mu_{Z_i})}_{\sum_{i=1}^{N}\sum_{k=1}^{K} Z_{ik} \log P(x_i \mid \mu_k, \Lambda_k) \;=\; \text{const} + \sum_{i=1}^{N} Z_{ij} \log P(x_i \mid \mu_j, \Lambda_j)}$$

$$+ \underbrace{\sum_{i=1}^{K} \log P(\Lambda_i)}_{\text{const} + \log P(\Lambda_j)} + \underbrace{\sum_{i=1}^{K} \log P(\mu_i \mid \Lambda_i)}_{\text{const} + \log P(\mu_j \mid \Lambda_j)}$$

مقادیر غیر از Λ_j constant

$$\longrightarrow P(x_i \mid \underline{\underline{Z_i}}, \overset{\mu \text{ و } \Lambda \text{ ثابت}}{\Lambda, \mu}) = \prod_{k=1}^{K} \left(P(x_i \mid \mu_k, \Lambda_k)\right)^{Z_{ik}}$$

$$Z_i = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \longrightarrow$$ sample iام از kامین component آمده.

$$E_{-j}[\log P(\text{all})] + \text{const} = \log P(\Lambda_j) + E_{\mu_j}[\log P(\mu_j \mid \Lambda_j)] + \underbrace{E\left[\sum_{i=1}^{N} Z_{ij} \log P(x_i \mid \mu_j, \Lambda_j)\right]}_{\sum_{i=1}^{N} \underbrace{E_{Z_{ij}}[Z_{ij}]}_{\alpha_i} E_{\mu_j}[\log P(x_i \mid \mu_j, \Lambda_j)]}$$

$$+ \text{const}\,2$$

$$\frac{|\Lambda_j|^{\frac{1}{2}}}{(2\pi)^{\frac{d}{2}}} \exp\left(\underbrace{-\tfrac{1}{2}(x_i-\mu_j)^T \Lambda_j (x_i-\mu_j)}_{\text{اسکالر}}\right)$$

$$\operatorname{tr}(\,''\,)$$

** با توجه به

$$-\tfrac{1}{2}\operatorname{tr}\left((x_i-\mu_j)(x_i-\mu_j)^T \Lambda_j\right)$$

$$-\tfrac{1}{2}\operatorname{tr}\left(\underbrace{E[(x_i-\mu_j)(x_i-\mu_j)^T]}_{A} \Lambda_j\right)$$

$$* \operatorname{tr}(ABC) = \operatorname{tr}(CAB) = \operatorname{tr}(BCA) \;**$$

$$A = \beta_0(\mu_k - m_0)(\mu_k - m_0)^T + W_0^{-1} + \sum_{n=1}^{N} r_{nk}(X_n - \mu_k)^T(X_n - \mu_k)$$
$$- \beta_k(\mu_k - m_k)^T(\mu_k - m_k) \; (*)$$

$$\frac{\nu_0 - D - 1 + \overbrace{\sum_{n=1}^{N} r_{nk}}^{N_k}}{2} \ln|\Lambda_k| - \tfrac{1}{2}\operatorname{Tr}[A \Lambda_k]$$

$$\log(\beta \Lambda_k) + \left(\frac{-(x-0)^2}{2(\beta\Lambda_k)^{-2}}\right)$$

$$= \log \underbrace{P(\Lambda_j)}_{\downarrow} + E_{\mu_j}[\log P(\mu_j \mid \Lambda_j)] + \sum_{i=1}^{n} \alpha_i \tfrac{1}{2} \operatorname{tr}(A \Lambda_j)$$

$\text{Wishart}(\Lambda_0, \nu)$

$$\overset{\exp(\;)}{\Longrightarrow} P(\Lambda_j \mid \text{other}) = \text{Wishart}(\Lambda_0, \nu) + \mathcal{N}(\mu_k \mid (\beta_k \Lambda_k)^{-1}) + \text{Wishart}(\Lambda', \nu')$$

$$\Lambda' \rightarrow \sum_{n=1}^{N} (X_n - \mu_k)(X_n - \mu_k)^T$$

$$\Longrightarrow P(\Lambda_j \mid \text{Other}) = \text{Wishart}\left((\Lambda_0^{-1} + \beta \mu_k \mu_k^T + \sum_{n=1}^{N} (X_n - \mu_k)(X_n - \mu_k)^T Z_n)^{-1}, \; \underbrace{\nu + 1 + \sum_{n=1}^{N} Z_{nk}}\right)$$

با توجه به (*)

$$\nu_k = \nu + 1 + \sum_{n=1}^{N} Z_{nk}$$

$$\log P(Z_{nk=1} \mid \text{Other}) = E_{-nk}\left[\log P(\text{all})\right] + \text{Const}$$

$$\log P(\text{all}) = \underbrace{\log P(\pi)}_{\text{const}} + \sum_{i=1}^{N} \log P(Z_i \mid \pi) + \sum_{i=1}^{N} \log P(X_i \mid Z_i, \Lambda_{Z_i}, \mu_{Z_i})$$

$$+ \underbrace{\sum_{i=1}^{K} \log P(\Lambda_i)}_{\text{Const}} + \underbrace{\sum_{i=1}^{K} \log P(\mu_i \mid \Lambda_i)}_{\text{const}}$$

$$= \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ik} \log \pi_k + \sum_{n=1}^{N} \sum_{k=1}^{K} \log \underbrace{N(X_n \mid \mu_k, \Lambda_k^{-1})}_{\frac{\Lambda_k}{(2\pi)^{\frac{D}{2}}} \exp\left(-\frac{1}{2}(x_i - \mu_j)^T \Lambda_j (x_i - \mu_j)\right)} \times Z_{nk}$$

$$\Rightarrow \log P(Z_{nk=1} \mid \text{Other}) = E_{\pi}\left[\log P(Z \mid \pi)\right] + E_{\mu, \Lambda}\left[\log P(X \mid Z, \mu, \Lambda)\right] + \text{const}$$

$$= \sum_{n=1}^{N} \sum_{k=1}^{K} \overset{1}{Z_{nk}} \left(E\left[\log \pi_k\right] + \frac{1}{2} E\left[\log \lambda_k\right] - \frac{D \log(2\pi)}{2} - \frac{1}{2} E_{\mu_k, \Lambda_k}\left[(x_n - \mu_k)^T \Lambda_k (x_n - \mu_k)\right] \right)$$

$$P(Z_{nk} = 1 \mid \text{Other}) = \prod_{n=1}^{N} \prod_{k=1}^{K} \underbrace{\exp\left(E\left[\log \pi_k\right] + \frac{1}{2} E\left[\log \lambda_k\right] - \frac{D}{2} \log(2\pi) - \frac{1}{2} E_{\mu_k, \Lambda_k}\left[(X_n - \mu_k)^T \Lambda_k (x_n - \mu_k)\right] \right)}_{\rho_{nk} \longrightarrow \pi_k N(X_n \mid \mu_k, \Lambda_k^{-1}) \; \xi_k}$$

$$P(Z_{nk} = 1 \mid \text{Other}) \propto \prod_{n=1}^{N} \prod_{k=1}^{K} \rho_{nk}^{\overset{1}{Z_{nk}}} = \prod_{n=1}^{N} \prod_{k=1}^{K} \frac{\rho_{nk}}{\sum_{i=1}^{K} \rho_{ni}} \quad r_{nk} \propto \pi_k |\Lambda_k|^{\frac{1}{2}} \exp\left(-\frac{1}{2}(x - \mu_k)^T \Lambda_k (x - \mu_k)\right)$$

$\Downarrow$

categorical distribution

$$\Rightarrow P(Z_{nk} = 1 \mid \text{other}) = \text{Categorical}\left(\frac{\xi_1}{\sum_{k=1}^{K} \xi_k}, \ldots, \frac{\xi_K}{\sum_{k=1}^{K} \xi_K} \right)$$

Scanned with CamScanner

$$\log P(\pi \mid \text{other}) = E_{-\pi}\left[\log P(\text{all})\right] + \text{const}$$

$$\log P(\text{all}) = \log P(\pi) + \log P(Z \mid \pi) + \underbrace{\sum_{i=1}^{N} \log P(X_i \mid Z_i, \Lambda_{z_i}, \mu_{z_i})}_{\text{const}}$$

$$+ \underbrace{\sum_{i=1}^{K} \log P(\Lambda_i)}_{\text{const}} + \underbrace{\sum_{i=1}^{K} \log P(\mu_i \mid \Lambda_i)}_{\text{const}}$$

$$\Rightarrow \log P(\pi \mid \text{other}) = \left[\underbrace{\log P(\pi)}_{\log C(\alpha_0) + \sum_{k=1}^{K} (\alpha_0 - 1)\log \pi_k}\right] + E_{-\pi}\left[\underbrace{\log P(Z \mid \pi)}_{\sum_{i=1}^{N}\sum_{k=1}^{K} Z_{ik} \log \pi_k}\right] + \text{const}$$

$$P(Z \mid \pi) = \prod_{n=1}^{N} \prod_{k=1}^{K} \pi_k^{z_{ik}}$$

$$\log P(\pi \mid \text{other}) = \underbrace{(\alpha_0 - 1) \sum_{k=1}^{K} \log \pi_k + \sum_{k=1}^{K} \sum_{i=1}^{N} \underbrace{E[Z_{ik}]}_{r_{ik}} \log \pi_k}_{\sum_{k=1}^{K} \log \pi_k \left[\alpha_0 - 1 + \sum_{i=1}^{N} r_{ik}\right]} + \text{Const}$$

$$\xrightarrow{e(\;)} P(\pi \mid \text{other}) = \prod_{k} \left(\pi_k^{\alpha_k - 1} \times \prod_{i=1}^{N} \pi_k \, \pi_k^{r_{ik}}\right) + \text{Const}$$

* باتوجه به

$$\Rightarrow P(\pi \mid \text{other}) = \text{Dirichlet}\left(\alpha_1 + \sum_{n=1}^{N} Z_{n1}, \ldots, \alpha_K + \sum_{n=1}^{N} Z_{nK}\right)$$

$$\boxed{\text{Dirichlet} = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_K)} \prod_{k=1}^{K} \pi^{\alpha_k - 1}} \; *$$

$$\alpha_0 = \sum_{k=1}^{K} \alpha_k$$

Scanned with CamScanner

$$\log P(\mu_K \mid \text{other}) = E_{-K}\left[\log P(\text{all})\right] + \text{const}$$

$$\log P(\text{all}) = \underbrace{\log P(\pi)}_{\text{const}} + \underbrace{\sum_{i=1}^{N} \log P(z_i \mid \pi)}_{\text{const}} + \overbrace{\sum_{i=1}^{N} \log P(x_i \mid z_i, \Lambda_{z_i}, \mu_{z_i})}^{\text{const} + \sum_{i=1}^{N} z_{ij} \log P(x_i \mid \mu_K, \Lambda_K)}$$

$$+ \underbrace{\sum_{i=1}^{K} \log P(\Lambda_j)}_{\text{const}} + \sum_{i=1}^{K} \underbrace{\log P(\mu_i \mid \Lambda_i)}_{\prod_{k=1}^{K} N(\mu_K \mid \mu_0, (\beta_0 \Lambda_K)^{-1})}$$

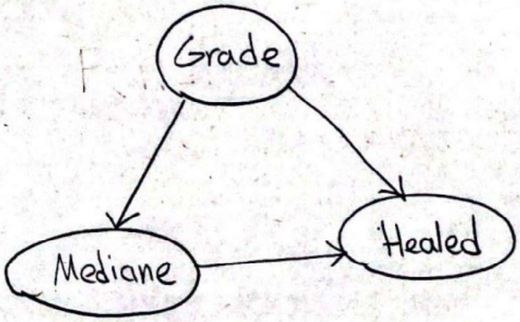
$$\log P(\text{all}) = \text{const} + \sum_{K=1}^{K}\left[\sum_{n=1}^{N} N(X_n \mid \mu_K, \Lambda^{-1}) Z_{nK} + N(\mu_K \mid \mu_0, (\beta_0 \Lambda_K)^{-1})\right]$$

$$\Rightarrow P(\mu_K \mid \text{Other}) = \exp\left(\sum_{K=1}^{K}\left[\sum_{n=1}^{N} \underbrace{N(X_n \mid \mu_K, \Lambda_K^{-1})}_{*} \times Z_{nK} + \underbrace{N(\mu_K \mid \mu_0, (\beta_0 \Lambda_K)^{-1})}_{**}\right]\right)$$

$$\propto \exp\left[\mu_K^{T}\left(\Lambda_K \sum_{n=1}^{N} X_n Z_{nK}\right) + \left(-\tfrac{1}{2}\mu_K^{T}\left(\beta \Lambda_K + \Lambda_K \sum_{n=1}^{N} Z_{nK}\right)\mu_K\right)\right]$$

Problem ④:

	medicine A	medicine B
G=1 fatty liver grade 1	$\frac{۸۱}{۸۷} = ۹۳\%$	$\frac{۲۳۴}{۲۷۰} = ۸۷\%$
G=4 " " 4	$\frac{۱۹۲}{۲۶۳} = ۷۳\%$	$\frac{۵۵}{۸۰} = ۶۹\%$
Both	$\frac{۸۱+۱۹۲}{۸۷+۲۶۳} = \frac{۲۷۳}{۳۵۰} = ۷۸\%$	$\frac{۵۵+۲۳۴}{۲۷۰+۸۰} = \frac{۲۸۹}{۳۵۰} = ۸۳\%$



$$P(M, H, G) = P(G)P(M|G)P(H|M, G)$$

$$P(X_1, \dots X_n | do x_i) = \prod_{i \neq j} p(X_j | PA_j(x_i))$$

$$P(\text{Healed} = \text{True} | do(\text{Med} = A)) = P(H=T | M=A, G=1)P(G=1) + P(H=T|M=A, G=4)P(G=4)$$

$$= \frac{۸۱}{۸۷} \times \frac{۸۷+۲۷۰}{۷۰۰} + \frac{۱۹۲}{۲۶۳} \times \frac{۲۶۳+۸۰}{۷۰۰} \simeq \boxed{۰/۸۳۲}$$

$$P(\text{Healed} = \text{True} | do(\text{Med} = B)) = P(H=T|M=B, G=1)P(G=1) + P(H=T|M=B, G=4)P(G=4)$$

$$= \frac{۲۳۴}{۲۷۰} \times \frac{۸۷+۲۷۰}{۷۰۰} + \frac{۵۵}{۸۰} \times \frac{۲۶۳+۸۰}{۷۰۰} \simeq \boxed{۰/۷۸۱۸}$$

* مریض‌هایی که grade 1 هستند داروی A را بیشتر از دسته دیگر دریافت می‌کنند.
* مریض‌هایی که داروی B را دریافت کردند، کمتر Heal شدند.

اگر هر دو گروه را باهم حساب کنیم نتیجه B بهتر است.

* با توجه به نتایج بهتر که داروی B دارد (۸۳٪) اما داروی A Healed rate بالاتری دارد. ۸۳/۲٪ در مقابل ۷۸/۱۸٪.
در نتیجه داروی A بهتر است.

Problem 5:

ابتدا کتابخانه های لازم را import کرده و داده های فایل csv را می خوانیم:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

```
df = pd.read_csv("ngaussian (1).csv" , delimiter=' ',header=None)
```

شکل ۱- کتابخانه های مورد نیاز و خواندن داده ها

[3]:

	0	1
0	-1.727048	-3.599960
1	-1.445891	-3.270109
2	-0.458869	-1.148513
3	1.317966	2.516560
4	1.153506	2.006051
...	...	...
995	2.068504	5.060544
996	-2.633527	-5.427452
997	1.200837	1.701804
998	-2.797796	-5.378348
999	-1.159075	-1.555699

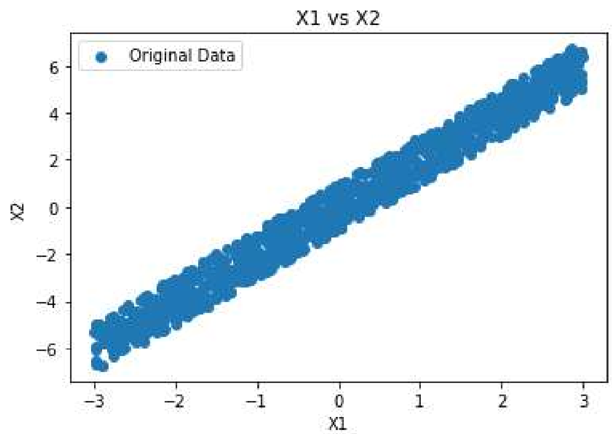
1000 rows × 2 columns

شکل ۲- نمایی از داده ها

نمایش داده ها به شکل زیر است:

```python
X1=df[0].values.reshape(-1, 1)
X2=df[1].values.reshape(-1, 1)
```

```python
plt.scatter(X1, X2 , label='Original Data')
plt.legend()
plt.xlabel('X1')
plt.ylabel('X2')
plt.title("X1 vs X2")
plt.show()
```



شکل ۳- نمایش داده ها

(a

در قسمت اول داده های X2 را روی X1 بصورت زیر فیت میکنیم و β_{12} را بدست می آوریم:

SCM M1

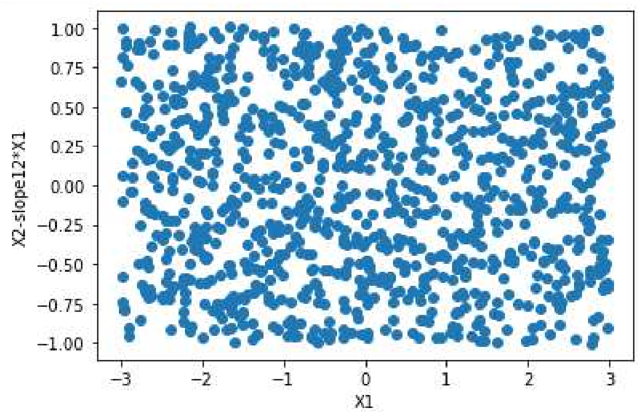
```python
model = LinearRegression(fit_intercept=False)

# Fit the model to the data
model.fit(X1, X2)

intercept = model.intercept_
slope12 = model.coef_[0]

print(f"B12 is: {slope12}")
```

```
B12 is: [2.00326517]
```

شکل ۴- فیت داده ها قسمت اول

β_{12} در این حالت ۲,۰۰۳۲۶۵۱۷ بدست آمد. در مرحله بعدی X_1 را برحسب $X_2 - \beta_{12}X_1$ رسم میکنیم:

```
[9]: plt.scatter(X1, X2-slope12*X1)
     plt.xlabel('X1')
     plt.ylabel('X2-slope12*X1')
     plt.show()
```



شکل ۵ - X1 برحسب X2 − β12X1

b) بصورت مشابه قسمت قبل داده های X1 را روی X2 بصورت زیر فیت میکنیم و β_{21} را بدست می آوریم:

SCM M2

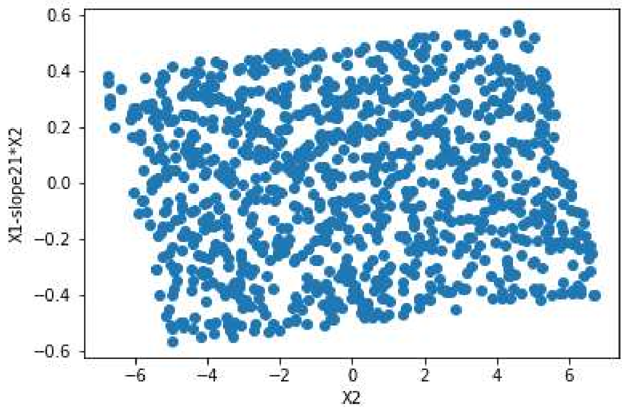
```
model = LinearRegression(fit_intercept=False)

# Fit the model to the data
model.fit(X2, X1)

intercept = model.intercept_
slope21 = model.coef_[0]

print(f"B21 is: {slope21}")
```

```
B21 is: [0.48538]
```

شکل ٦ - فیت داده ها قسمت دوم

β_{21} در این حالت ۰,۴۸۵۳۸ بدست آمد(عکس حالت قبل). در مرحله بعدی X_2 را برحسب $X_1 - \beta_{21}X_2$ رسم میکنیم:

```
[12]: plt.scatter(X2, X1-slope21*X2)
      plt.xlabel('X2')
      plt.ylabel('X1-slope21*X2')
      plt.show()
```



شکل ۷ - X_2 برحسب $X_1 - \beta_{21}X_2$

(c

مدل با توزیع یکنواخت تر نمودار residual تناسب بهتری با داده ها فراهم می کند. یک نمودار residual توزیع شده به طور مساوی در نمودار اول نشان می دهد که مدل به طور دقیق رابطه بین

X1 و X2 - β12 * X1 را توضیح می دهد. این نشان می‌دهد که X1 علت X2 است، زیرا مدل به دقت تغییرات X2 را بر اساس X1 ثبت می‌کند. به عبارت دیگر، X1 پیش بینی بهتری برای X2 نسبت به عکس آن است.

توزیع ناهموار نمودار دوم نشان می دهد که مدل در نشان دادن رابطه بین X2 و X1 - β21 * X2 به اندازه کافی مؤثر نیست.

- رویکرد خطی-گاوسی معمولاً تنها به مجموعه‌ای از مدل‌های ممکن منتهی می‌شود که در ساختار همبستگی شرطی آنها معادل است.
- یک تنظیم خطی غیر گاوسی اجازه می دهد تا مدل علی کامل بدون پارامترهای نامشخص تخمین زده شود.