Assignment Three

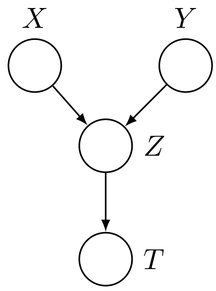
Matthew C. Scicluna
Département d'Informatique et de Recherche Opérationnelle
Université de Montréal
Montréal, QC H3T 1J4
matthew.scicluna@umontreal.ca

October 29, 2017

1 DGM

Given the following DGM G the implied factorization for any joint $p \in \mathcal{L}(G)$ is

$$p(X, Y, Z, T) = f_X(X) f_Y(Y) f_Z(Z; X, Y) f_T(T; Z)$$



It is not true that for any $p \in \mathcal{L}(G)$ $X \perp Y \mid T$. For a counterexample, take $X, Y \sim Bern\left(\frac{1}{2}\right)$, with $T = Z = X + Y$. Then,

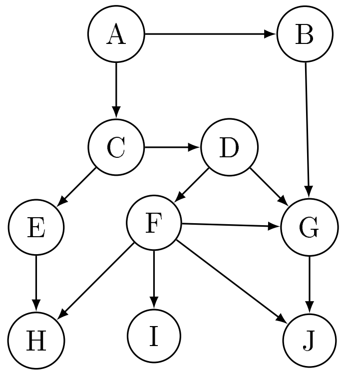
$$P(X = 1, Y = 0 \mid Z = 1) = \frac{1}{2}$$

but

$$P(X = 1 \mid Z = 1) = P(Y = 0 \mid Z = 1) = \frac{1}{2}$$

Hence $X \not\perp Y \mid T$, for this $p \in \mathcal{L}(G)$, as needed.

2 d-Separation in DGM



(a) FALSE, consider the path $(C, A), (A, B)$

(b) TRUE

(c) FALSE, consider the path $(C, D), (D, G), (G, B)$

(d) TRUE

(e) FALSE, consider the path $(C, A), (A, B), (B, G)$

(f) FALSE, consider the path $(C, D), (D, G)$

(g) TRUE

(h) FALSE, consider the path $(C, E), (E, H), (H, F), (F, G)$

(i) TRUE

(j) FALSE, consider the path $(B, A), (A, C), (C, D), (D, F), (F, I)$

3 Positive interactions in-V-structure

Given X, Y, Z binary RV's with a joint distribution parameterized by $X \to Z \leftarrow Y$, with $a = P(X = 1), b = P(X = 1|Z = 1), c = P(X = 1|Z = 1, Y = 1)$. We notice that if we set $X \sim Bern\left(\frac{1}{2}\right)$, $Y \sim Bern\left(\frac{1}{2}\right)$ and if we denote $\alpha = P(Z = 1 \mid X = 1, Y = 1), \beta = P(Z = 1 \mid X = 1, Y = 0), \gamma = P(Z = 1 \mid X = 0, Y = 1)$ and $\delta = P(Z = 1 \mid X = 0, Y = 0)$ that we have, by cancellation of marginal probabilities of X and Y, that $c - b = \frac{\alpha}{\alpha+\gamma} - \frac{\alpha+\beta}{\alpha+\beta+\gamma+\delta}$. We then set $\alpha, \beta, \gamma, \delta$ accordingly to satisfy the inequality relations between a, b and c with fixed $a = \frac{1}{2}$.

(a) (i) $X \sim Bern\left(\frac{1}{2}\right)$, $Y \sim Bern\left(\frac{1}{2}\right)$ and $Z = 1 - X \oplus Y$. Then $a = \frac{1}{2}$ but $c = 0$

(ii) Again, we fix $X \sim Bern\left(\frac{1}{2}\right)$, $Y \sim Bern\left(\frac{1}{2}\right)$ and Z has the following probability table:

X	Y	P(Z=1)
0	0	0.1
0	1	0.8
1	0	1
1	1	1

Then $a = \frac{1}{2}$, $c = \frac{\frac{1}{2}}{0.8+0.2\frac{1}{2}} = \frac{5}{9}$ and $b = \frac{\frac{1}{2}}{0.83+0.17\frac{1}{2}} \approx 0.855$, and $a < c < b$.

(iii) Again, we fix $X \sim Bern\left(\frac{1}{2}\right)$, $Y \sim Bern\left(\frac{1}{2}\right)$ and Z has the following probability table:

X	Y	P(Z=1)
0	0	1
0	1	0.8
1	0	0
1	1	1

Then $a = \frac{1}{2}$, $c = \frac{1}{1.8}$ and $b = \frac{1}{2.8}$, and $b < a < c$.

(b) (i) The semantic here is that Z is a negated XOR gate for X and Y. Hence, knowing that both Y and Z are "on" means that X must not be.

(ii) Semantically, Z will certainly be "on" if X is, and probably will be "on" if Y is. So Z being "on" gives evidence for X, but having Y "on" as well can explain away Z (i.e. it is more likely that Z was caused by only Y).

(iii) Semantically, Z is likely to be "on" unless X is "on" and Y isn't. So if Z is "on", X is less likely to be, since Y would have to be "on" too.

4 Flipping a covered edge in a DGM

Let $G = (V, E)$ be a DAG. We say that a directed edge $(i, j) \in E$ is a covered edge if and only if $\pi_j = \pi_i \cup \{i\}$. Let $G' = (V, E')$, with $E' = (E - \{(i, j)\}) \cup \{(j, i)\}$. Prove that $\mathcal{L}(G) = \mathcal{L}(G')$.

PROOF: Fix i, j, G and E. Denote π'_k as the set of parents for a node k under E'. We note that $\pi'_k = \pi_k$ for $k \neq i, j$ and $\pi'_i = \pi_i \cup \{j\}$ and $\pi'_j = \pi_i$.
For the forward direction we let $p \in \mathcal{L}(G)$. We want to show that $p(x_v) = \prod_{k=1}^{n} p(x_k|x_{\pi'_k})$
Notice that:

$$p(x_v) = \prod_{k \neq i,j} p(x_k|x_{\pi_k}) P(x_i|x_{\pi_i}) P(x_j|x_{\pi_i}, x_i) \tag{4.1}$$

$$= \prod_{k \neq i,j} p(x_k|x_{\pi_k}) P(x_i, x_j|x_{\pi_i}) \tag{4.2}$$

$$= \prod_{k \neq i,j} p(x_k|x_{\pi_k}) P(x_j|x_{\pi_i}) P(x_i|x_{\pi_i}, x_j) \tag{4.3}$$

$$= \prod_{k \neq i,j} p(x_k|x_{\pi'_k}) P(x_j|x_{\pi'_j}) P(x_i|x_{\pi'_i}) \tag{4.4}$$

$$= \prod_{k=1}^{n} p(x_k|x_{\pi'_k}) \tag{4.5}$$

As needed. For the reverse direction, let $p \in \mathcal{L}(G')$

$$
\begin{aligned}
p(x_v) &= \prod_{k \neq i,j} p(x_k|x_{\pi'_k}) P(x_i|x_{\pi'_j}, x_j) P(x_j|x_{\pi'_j}, x_i) && (4.6) \\
&= \prod_{k \neq i,j} p(x_k|x_{\pi'_k}) P(x_i|x_{\pi_i}) P(x_j|x_{\pi_i}, x_i) && (4.7) \\
&= \prod_{k \neq i,j} p(x_k|x_{\pi_k}) P(x_i|x_{\pi_i}) P(x_j|x_{\pi_j}) && (4.8)
\end{aligned}
$$

And (4.5) and (4.8) complete the proof.

5 Equivalence of directed tree DGM with undirected tree UGM

Let G be a directed tree and G' be its corresponding undirected tree. Prove that $\mathcal{L}(G) = \mathcal{L}(G')$.

PROOF: For the forward direction we set the edge potentials $\psi_{(i,j)}(x_i, x_j) = P(x_j|x_i)$. For the node potentials we set $\psi_r(x_r) = P(x_r)$, where x_r is the root of the tree, and 1 for the rest. It is clear that

$$
P(x_v) = \prod_{i \in V} P(x_i|x_{\pi_i}) = \psi_r(x_r) \prod_{(\pi_i, i) \in E} \psi_{(\pi_i, i)}(x_{\pi_i}, x_i)
$$

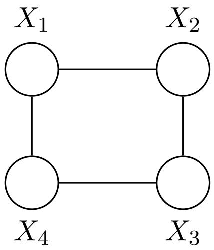
We note that $Z = 1$ since

$$
\begin{aligned}
Z &= \sum_{x_v} \psi_r(x_r) \prod_{(\pi_i, i) \in E} \psi_{(\pi_i, i)}(x_{\pi_i}, x_i) \\
&= \sum_{x_r} \psi_r(x_r) \prod_{(\pi_i, i) \in E} \sum_{x_i} \psi_{(\pi_i, i)}(x_{\pi_i}, x_i) \\
&= \sum_{x_r} P(x_r) \prod_{(\pi_i, i) \in E} \sum_{x_i} P(x_i|x_{\pi_i}) = 1
\end{aligned}
$$

For the reverse direction, TBD

6 Hammersley-Clifford Counter example

Given $P(0,0,0,0) = P(1,0,0,0) = P(1,1,0,0) = P(1,1,1,0) = P(0,0,0,1) = P(0,0,1,1) = P(0,1,1,1) = P(1,1,1,1) = \frac{1}{8}$ and the following graph:



We want to show that $p \notin \mathcal{L}(G)$.

PROOF: Suppose $p \in \mathcal{L}(G)$, then there are ψ potentials such that

$$P(x_1, x_2, x_3, x_4) = \psi_{x_1,x_2}(x_1, x_2)\psi_{x_2,x_3}(x_2, x_3)\psi_{x_3,x_4}(x_3, x_4)\psi_{x_4,x_1}(x_4, x_1)$$

Notice that $P(0,1,0,0) = 0$ implies that at least one of the following must be 0: $\psi_{x_1,x_2}(0,1)$, $\psi_{x_2,x_3}(1,0)$, $\psi_{x_3,x_4}(0,0)$, $\psi_{x_4,x_1}(0,0)$

We notice that if $\psi_{x_1,x_2}(0,1) = 0$ then $P(0,1,1,1) = 0$ which contradicts that $P(0,1,1,1) = \frac{1}{8}$. Similarily, if $\psi_{x_2,x_3}(1,0) = 0$ then $P(1,1,0,0) = 0$, contradicting that it is $\frac{1}{8}$. The same reasoning shows why $\psi_{x_3,x_4}(0,0) \neq 0$. Finally, if $\psi_{x_4,x_1}(0,0) = 0$ then $P(0,0,0,0) = 0$, which again is a contradiction.