

Solved exercises for 'Probabilistic Graphical Models: Principles and techniques'

Gianmarco Cafferata

November 28, 2020

2 Foundations

2.1 Exercise 2.1

2.1.1 (a) $P(\emptyset) = 0$

From the axioms and set theory:

$$\emptyset, \Omega \in \mathcal{S}, \emptyset \cap \Omega = \emptyset \Rightarrow P(\emptyset \cup \Omega) = P(\emptyset) + P(\Omega) \quad (1)$$

$$P(\Omega) = 1 \quad (2)$$

$$\emptyset \cup \Omega = \Omega \quad (3)$$

With (2) and (3) we got

$$P(\Omega) = P(\emptyset \cup \Omega) = 1 \Rightarrow^{(1)} P(\emptyset \cup \Omega) = P(\emptyset) + P(\Omega) = 1 \Rightarrow P(\emptyset) = 1 - P(\Omega) =^{(2)} 0$$

2.1.2 (b) If $\alpha \subseteq \beta$, then $P(\alpha) \leq P(\beta)$

Given the special case that $\alpha = \beta \Rightarrow P(\alpha) = P(\beta)$ and the inequality is directly verified.

Lets analyze the other case where $\alpha \subset \beta$ but $\alpha \neq \beta$, from the axioms and set theory we got:

Let θ be $\theta = \beta - \alpha$

$$\beta = \alpha \cup \theta \quad (1)$$

$$\alpha, \theta \in \mathcal{S}, \alpha \cap \theta = \emptyset \Rightarrow P(\alpha \cup \theta) = P(\alpha) + P(\theta) \quad (2)$$

Then $P(\beta) =^{(1)} P(\alpha \cup \theta) =^{(2)} P(\alpha) + P(\theta)$, but because $\theta \in \mathcal{S}$ we know $P(\theta) \geq 0$, therefore we reach $P(\beta) = P(\alpha) + P(\theta) \geq P(\alpha)$.

2.1.3 (c) $P(\alpha \cup \beta) = P(\alpha) + P(\beta) - P(\alpha \cap \beta)$

Let $\alpha = \alpha' \cup \theta$ and $\beta = \beta' \cup \theta$ such that $\alpha' \cap \theta = \emptyset$ and $\beta' \cap \theta = \emptyset$.

Then $\alpha \in \mathcal{S} \Rightarrow \alpha' \in \mathcal{S} \wedge \theta \in \mathcal{S}$ and the same goes for beta.

Because of set theory and the axioms we know:

$$\alpha', \theta \in \mathcal{S}, \alpha' \cap \theta = \emptyset \Rightarrow P(\alpha' \cup \theta) = P(\alpha) = P(\alpha') + P(\theta) \quad (1)$$

$$\beta', \theta \in \mathcal{S}, \beta' \cap \theta = \emptyset \Rightarrow P(\beta' \cup \theta) = P(\beta) = P(\beta') + P(\theta) \quad (2)$$

$$\alpha \cup \beta = \alpha' \cup \beta' \cup \theta \quad (3)$$

Because $(\alpha' \cup \beta') \cap \theta = \emptyset$ and $\alpha' \cap \beta' = \emptyset$ then:

$$P(\alpha' \cup \beta' \cup \theta) = P(\alpha') + P(\beta') + P(\theta) \quad (4)$$

And finally the proof:

$$\begin{aligned}
P(\alpha \cup \theta) & \stackrel{(3)}{=} P(\alpha' \cup \beta' \cup \theta) \\
P(\alpha \cup \theta) & \stackrel{(4)}{=} P(\alpha') + P(\beta') + P(\theta) \iff \text{add } P(\theta) \text{ on both sides} \\
P(\alpha \cup \theta) + P(\theta) & = P(\alpha') + P(\beta') + 2P(\theta) \iff P(\alpha') + P(\theta) = P(\alpha) \\
P(\alpha \cup \theta) + P(\theta) & = P(\alpha) + P(\beta') + P(\theta) \iff P(\beta') + P(\theta) = P(\beta) \\
P(\alpha \cup \theta) + P(\theta) & = P(\alpha) + P(\beta) \iff \theta = \alpha \cap \beta \\
P(\alpha \cup \theta) & = P(\alpha) + P(\beta) - P(\alpha \cap \beta)
\end{aligned}$$

2.2 Exercise 2.2

2.2.1 (a) Show that for binary random variables X, Y the event-level independence $(x^0 \perp y^0)$ implies random variable independence

		Y		
		0	1	
X	0	$P(X=0, Y=0)$	$P(X=0, Y=1)$	$P(X=0)$
	1	$P(X=1, Y=0)$	$P(X=1, Y=1)$	$P(X=1)$
		$P(Y=0)$	$P(Y=1)$	

We know that $x^0 \perp y^0 \iff y^0 \perp x^0$ so we need to prove $x^1 \perp y^0, x^0 \perp y^1$ and $x^1 \perp y^1$. Lets start with $x^1 \perp y^0$:

$$\begin{aligned}
P(X=1|Y=0) & = 1 - P(X=0|Y=0) \iff x^0 \perp y^0 \\
P(X=1|Y=0) & = 1 - P(X=0) \iff \text{since } \sum_{i=0}^1 P(X=x^i) = 1 \text{ then } 1 - P(X=0) = P(X=1)
\end{aligned}$$

$$P(X=1|Y=0) = P(X=1) \quad (1)$$

$$P(X=1 \cap Y=0) = P(X=1|Y=0)P(Y=0) \stackrel{(1)}{=} P(X=1)P(Y=0) \Rightarrow x^1 \perp y^0$$

The same goes for $y^1 \perp x^0$. Now we prove $x^1 \perp y^1$:

$$P(X=1|Y=1) = 1 - P(X=0|Y=1) \stackrel{X=0 \perp Y=1}{=} 1 - P(X=0) = P(X=1) \quad (2)$$

$$P(X=1 \cap Y=1) = P(X=1|Y=1)P(Y=1) \stackrel{(2)}{=} P(X=1)P(Y=1) \Rightarrow x^1 \perp y^1$$

We have proven $X = x \perp Y = y$ for all values $x \in \text{Val}(X), y \in \text{Val}(Y)$ therefore $X \perp Y$.

2.2.2 (b) Show a counterexample for nonbinary variables

		Y		
		0	1	
X	0	0.1	0.1	0.2
	1	0.1	0.4	0.5
	2	0.3	0.0	0.3
		0.5	0.5	

Here we can see that $P(X = 0|Y = 0) = \frac{P(X=0,Y=0)}{P(Y=0)} = \frac{0.1}{0.5} = 0.2 = P(X = 0)$ therefore $X = 0 \perp Y = 0$. Les see what happens for $P(X = 1|Y = 1)$:

$$P(X = 1|Y = 1) = \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{0.4}{0.5} = 0.8 \neq P(X = 1) = 0.5$$

We have found one tuple $x \in Val(X), y \in Val(Y)$ where $X = x \not\perp Y = y$ therefore $X \not\perp Y$.

2.2.3 (c) Is it the case that, for a binary-valued variable Z we have that $(X \perp Y|z^0)$ implies $(X \perp Y|Z)$?

No, here is a counterexample:

		Y		
		0	1	
X	0	0.15	0.15	0.3
	1	0.15	0.15	0.3
		0.3	0.3	

Table 1: Joint distribution for $z=0$, with $P(z = 0) = 0.6$

		Y		
		0	1	
X	0	0.1	0.2	0.3
	1	0.1	0.0	0.1
		0.2	0.2	

Table 2: Joint distribution for $z=1$, with $P(z = 1) = 0.4$

Lets find $P(X = x \cap Y = y|z^0)$:

$$P(X = x \cap Y = y|z^0) = \frac{P(X = x, Y = y, z = 0)}{P(z = 0)} = \frac{0.15}{0.6} = 0.25 \quad (1)$$

Lets remember $(X \perp Y|z^0) \iff P(X = x \cap Y = y|z^0) = P(X = x|z^0) * P(Y = y|z^0)$:

$$P(X = x|z^0) * P(Y = y|z^0) = \frac{0.3}{0.6} * \frac{0.3}{0.6} = 0.25 \stackrel{(1)}{=} P(X = x \cap Y = y|z^0)$$

So this example satisfies the hypothesis that $(X \perp Y|z^0)$, but for example:

$$P(X = 0 \cap Y = 1|z^1) = \frac{P(X = 0, Y = 1, Z = 1)}{P(Z = 1)} = \frac{0.2}{0.4} = 0.5 \quad (2)$$

$$P(X = 0|z^1) * P(Y = 1|z^1) = \frac{0.3}{0.4} * \frac{0.2}{0.4} = \frac{3}{8} \quad (3)$$

We know that $(X \perp Y|Z) \iff (X = x \perp Y = y|Z = z)$ for every $x \in \text{Val}(X), y \in \text{Val}(Y), z \in \text{Val}(Z)$. Also $(X = 0 \perp Y = 1|Z = 1) \iff P(X = 0 \cap Y = 1|z^1) = P(X = 0|z^1) * P(Y = 1|z^1)$ but $P(X = 0 \cap Y = 1|z^1) = 0.5 \neq P(X = 0|z^1) * P(Y = 1|z^1) = \frac{3}{8}$ therefore $(X = 0 \not\perp Y = 1|Z = 1)$.

I have found a case where $(X = 0 \not\perp Y = 1|Z = 1) \Rightarrow (X \not\perp Y|Z)$ given $(X \perp Y|z^0)$.

2.3 Exercise 2.3

2.3.1 $\alpha \cap \beta$

$$P(\alpha \cap \beta) = P(\alpha|\beta) * P(\beta) \quad (1)$$

Lets find the minimum, if $\alpha \cap \beta = \emptyset$:

$$P(\alpha \cap \beta) = 0$$

Since $P(a) \geq 0$ for all $a \in \mathcal{S}$ then $P(\alpha \cap \beta) < 0$ is not possible therefore $P(\alpha \cap \beta) \geq 0$. Lets find the maximum, particularly if $\alpha = \beta$:

$$P(\alpha \cap \beta) = P(\alpha) = P(\beta)$$

Since $0 \leq P(\alpha|\beta) \leq 1$ then $P(\alpha \cap \beta) \stackrel{(1)}{=} P(\alpha|\beta) * P(\beta) \leq P(\beta)$ therefore the particular case when $\alpha = \beta$ produces the maximum then $P(\alpha \cap \beta) \leq P(\alpha) = P(\beta)$.

2.3.2 $\alpha \cup \beta$

We have proven in exercise 2.1 that:

$$P(\alpha \cup \beta) = P(\alpha) + P(\beta) - P(\alpha \cap \beta) \quad (1)$$

$P(\alpha)$ and $P(\beta)$ are known, we can only minimize and maximize $P(\alpha \cap \beta)$.

The minimum occurs when $\alpha = \beta$ and $P(\alpha \cap \beta) = P(\alpha) = P(\beta)$.

The maximum occurs when $\alpha \cap \beta = \emptyset$.

2.4 Exercise 2.4

2.4.1 $P(\cdot|\alpha) \geq 0$

$$P(\cdot|\alpha) = \frac{P(\cdot \cap \alpha)}{P(\alpha)}$$

But we know $\cdot \cap \alpha$ is in the event space then because of the same definition $P(\cdot \cap \alpha) \geq 0$ and we know that $P(\alpha) > 0$. So $P(\cdot|\alpha) \geq 0$.

2.4.2 $P(\Omega|\alpha) = 1$

$$P(\Omega|\alpha) = \frac{P(\Omega \cap \alpha)}{P(\alpha)} =_{\Omega \cap \alpha = \alpha} \frac{P(\alpha)}{P(\alpha)} = 1$$

2.4.3 $\beta, \gamma \in \mathcal{S}, \beta \cap \gamma = \emptyset \Rightarrow P(\beta \cup \gamma|\alpha) = P(\beta|\alpha) + P(\gamma|\alpha)$

$$P(\beta \cup \gamma|\alpha) = \frac{P((\beta \cap \alpha) \cup (\gamma \cap \alpha))}{P(\alpha)} \quad (1)$$

Because of set theory:

$$\beta \cap \alpha \subset \beta, \gamma \cap \alpha \subset \gamma \quad (2)$$

$$\beta, \gamma, \alpha \in \mathcal{S} \Rightarrow (\beta \cap \alpha) \in \mathcal{S}, (\gamma \cap \alpha) \in \mathcal{S}, \quad (3)$$

$$(\beta \cap \alpha) \cap (\gamma \cap \alpha) \subset_{(2)} \beta \cap \alpha = \emptyset \Rightarrow (\beta \cap \alpha) \cap (\gamma \cap \alpha) = \emptyset \quad (4)$$

Then with (3) and (4) in (1):

$$\frac{P((\beta \cap \alpha) \cup (\gamma \cap \alpha))}{P(\alpha)} =_{(3),(4)} \frac{P(\beta \cap \alpha)}{P(\alpha)} + \frac{P(\gamma \cap \alpha)}{P(\alpha)} = P(\beta|\alpha) + P(\gamma|\alpha)$$

2.5 Exercise 2.5

We say that P satisfies $(\mathbf{X} \perp \mathbf{Y}|\mathbf{Z})$ when $(\mathbf{X} = \mathbf{x} \perp \mathbf{Y} = \mathbf{y}|\mathbf{Z} = \mathbf{z})$. for all $\mathbf{x} \in \text{Val}(\mathbf{X}), \mathbf{y} \in \text{Val}(\mathbf{Y}), \mathbf{z} \in \text{Val}(\mathbf{Z})$

We also know that $(\mathbf{X} = \mathbf{x} \perp \mathbf{Y} = \mathbf{y}|\mathbf{Z} = \mathbf{z})$ if and only if $P(X = x \cup Y = y|Z = z) = P(X = x|Z = z)P(Y = y|Z = z)$

$$\mathbf{2.5.1} \quad P(X, Y|Z) = P(X|Z)P(Y|Z) \iff (\mathbf{X} \perp \mathbf{Y}|Z)$$

For all $\mathbf{x} \in Val(\mathbf{X}), \mathbf{y} \in Val(\mathbf{Y}), \mathbf{z} \in Val(\mathbf{Z})$:

$$\begin{aligned} P(X = x \cup Y = y|Z = z) &=_{\text{condition}} \\ P(X = x|Z = z)P(Y = y|Z = z) &\iff_{\text{event conditional independence}} \\ & (X = x \perp Y = y|Z = z) \end{aligned}$$

Above we have that:

$$(X = x \perp Y = y|Z = z) \text{ for all } \mathbf{x} \in Val(\mathbf{X}), \mathbf{y} \in Val(\mathbf{Y}), \mathbf{z} \in Val(\mathbf{Z})$$

Which is the definition of $(\mathbf{X} \perp \mathbf{Y}|Z)$.

$$\mathbf{2.5.2} \quad (\mathbf{X} \perp \mathbf{Y}|Z) \Rightarrow P(\chi) = \phi_1(\mathbf{X}, Z)\phi_2(\mathbf{Y}, Z)$$

$$\begin{aligned} P(X, Y|Z) &= \frac{P(X, Y, Z)}{P(Z)} = \frac{P(\chi)}{P(Z)} \Rightarrow P(\chi) = P(X, Y|Z) * P(Z) \\ P(\chi) &= P(X, Y|Z) * P(Z) = P(X|Z)P(Y|Z)P(Z) \end{aligned}$$

Let $\phi_1(X, Z) = P(X|Z)$ and $\phi_2(Y, Z) = P(Y|Z)P(Z)$, then:

$$P(\chi) = \phi_1(X, Z)\phi_2(Y, Z)$$

$$\mathbf{2.5.3} \quad P(\chi) = \phi_1(\mathbf{X}, Z)\phi_2(\mathbf{Y}, Z) \Rightarrow (\mathbf{X} \perp \mathbf{Y}|Z)$$

We have that:

$$P(Z) = \sum_x \sum_y P(\chi) = \sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z) \quad (1)$$

$$P(X, Y|Z) = \frac{P(\chi)}{P(Z)} =_{(1)} \frac{\phi_1(X, Z)\phi_2(Y, Z)}{\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z)} \quad (2)$$

$$P(X|Z) = \frac{P(X, Z)}{P(Z)} = \frac{\sum_y P(\chi)}{P(Z)} =_{(1)} \frac{\phi_1(X, Z) \sum_y \phi_2(Y, Z)}{\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z)} \quad (3)$$

$$P(Y|Z) = \frac{P(Y, Z)}{P(Z)} = \frac{\sum_x P(\chi)}{P(Z)} =_{(1)} \frac{\phi_2(Y, Z) \sum_x \phi_1(X, Z)}{\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z)} \quad (4)$$

And then multiplying (3) and (4):

$$\begin{aligned}
P(X|Z)P(Y|Z) &= \frac{\phi_1(X, Z)\phi_2(Y, Z)(\sum_x \phi_1(X, Z))(\sum_y \phi_2(Y, Z))}{(\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z))^2} \\
&= \frac{\phi_1(X, Z)\phi_2(Y, Z)(\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z))}{(\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z))^2} \\
&= \frac{\phi_1(X, Z)\phi_2(Y, Z)}{\sum_x \sum_y \phi_1(X, Z)\phi_2(Y, Z)} \stackrel{(2)}{=} P(X, Y|Z) \Rightarrow_{2.5.1} (\mathbf{X} \perp \mathbf{Y}|\mathbf{Z})^1
\end{aligned}$$

2.6 Exercise 2.6

$$P(X|Y) = \frac{P(X, Y)}{P(Y)} = \frac{\sum_z P(X, Y, Z = z)}{P(Y)} \quad (1)$$

$$\sum_z P(X, Z = z|Y) = \frac{\sum_z P(X, Y, Z = z)}{P(Y)} \quad (2)$$

And there we have that (1) is equal to (2)²

2.7 Exercise 2.7

¹This could be also an if and only if proof so 2.5.2 wouldn't be unnecessary

²Actually i don't know if i can use the sum over the joint distribution for (1) but the exercise was boring