MS Data Science

Probabilistic Graphical Models

Homework 1

Firas Jarboui
Imad El Hanafi
Promotion 2018

Professeur :
M. Umut Şimşekli

05 October 2017

Table des matières

1 Question 1

Let denote by $\mathcal{N}(x;\mu,\sigma^2)$ the gaussian distribution of mean μ and variance σ.

$$\mathcal{N}(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)$$

Let λ and $\mathcal{D}$ be two two random variables such that $p(\lambda) = \mathcal{N}(\lambda;0,\sigma_\lambda^2)$ and $p(\mathcal{D}/\lambda) = \mathcal{N}(\mathcal{D};\lambda,\sigma_{\mathcal{D}}^2)$. Using the Bayes' theorem we write

$$p(\lambda/\mathcal{D}) = \frac{p(\mathcal{D}/\lambda)p(\lambda)}{p(\mathcal{D})} \propto p(\mathcal{D}/\lambda)p(\lambda)$$

knowing the expression of Gaussian distribution, we get

$$\begin{aligned}
p(\lambda/\mathcal{D}) &\propto \exp\Big(-\frac{(\mathcal{D}-\lambda)^2}{2\sigma_{\mathcal{D}}^2}\Big)\exp\Big(-\frac{(\lambda)^2}{2\sigma_\lambda^2}\Big) \\
&\propto \exp\Big(-\frac{\sigma_\lambda^2(\mathcal{D}-\lambda)^2 + \sigma_{\mathcal{D}}^2\lambda^2}{2\sigma_\lambda^2\sigma_{\mathcal{D}}^2}\Big) \\
&\propto \exp\Big(-\frac{(\sigma_\lambda^2+\sigma_{\mathcal{D}}^2)\lambda^2 - 2\sigma_\lambda^2\lambda\mathcal{D} + \sigma_\lambda^2\mathcal{D}^2}{2\sigma_\lambda^2\sigma_{\mathcal{D}}^2}\Big) \\
&\propto \exp\Big(-\frac{1}{2\frac{\sigma_\lambda^2\sigma_{\mathcal{D}}^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2}}(\lambda^2 - 2\frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2}\lambda\mathcal{D} + \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2}\mathcal{D}^2)\Big)
\end{aligned}$$

Let

$$\sigma_1^2 = \frac{\sigma_\lambda^2\sigma_{\mathcal{D}}^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2}$$

And given that

$$\lambda^2 - 2\frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2}\lambda\mathcal{D} + \frac{\sigma_\lambda^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2}\mathcal{D}^2 = \lambda^2 - 2\frac{\sigma_1^2}{\sigma_{\mathcal{D}}^2}\lambda\mathcal{D} + \frac{\sigma_1^2}{\sigma_{\mathcal{D}}^2}\mathcal{D}^2 =^+ (\lambda - \frac{\sigma_1}{\sigma_{\mathcal{D}}}\mathcal{D})^2$$

Therefore

$$p(\lambda/\mathcal{D}) \propto \exp\Big(-\frac{(\lambda - \frac{\sigma_1}{\sigma_{\mathcal{D}}}\mathcal{D})^2}{2\sigma_1^2}\Big) \propto \mathcal{N}(\lambda; m_1, \sigma_1^2)$$

We conclude that $\lambda/\mathcal{D}$ follows (proportional) a normal distribution with mean

$$m_1 = \frac{\sigma_1}{\sigma_{\mathcal{D}}}\mathcal{D}$$

and variance

$$\sigma_1^2 = \frac{\sigma_\lambda^2\sigma_{\mathcal{D}}^2}{\sigma_\lambda^2+\sigma_{\mathcal{D}}^2} = \frac{1}{\frac{1}{\sigma_{\mathcal{D}}^2}+\frac{1}{\sigma_\lambda^2}}$$

2 Question 2

let's define the variable of this problem as follows :

— P : printing success (observable state)
— A : driver corruption
— B : software problem
— C : unplugged printer
— D : being out of paper
— E : being out of power
— L : fact that the lights are off (observable state)

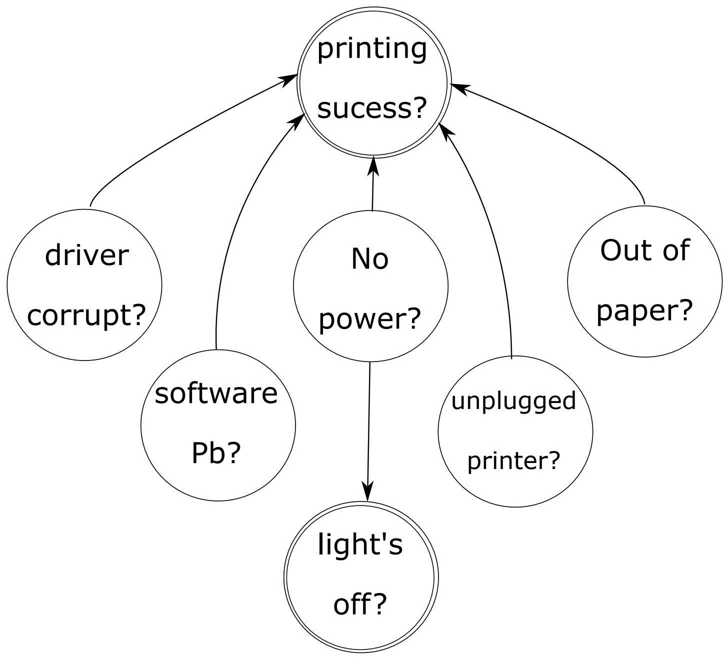

Figure 1 : Directed Graph for the printer's troubleshooter

3 Question 3

Directed graph

We use the following factorization to define the directed graph

$$p(A, B, D, F, T, L, M, X) = p(F|T, L)p(M)p(T|A)p(B|M)p(X|F)p(L|M)p(D|F, B)p(A)$$

Alternative representation

Undirected graph

We define the following set of clique functions :

$$\begin{cases} \mathcal{F}(A) = & p(A) \\ \mathcal{F}(A,T) = & p(T|A) \\ \mathcal{F}(T,F,L) = & p(F|T,L) \\ \mathcal{F}(F,X) = & p(X|F) \\ \mathcal{F}(F,D,B) = & p(D|F,B) \\ \mathcal{F}(B,M) = & p(B|M) \\ \mathcal{F}(M) = & p(M) \\ \mathcal{F}(M,L) = & p(L|M) \end{cases} \tag{3.1}$$

(A), (A, T), (T, F, L), (F, X), (F, B, D),(B, M),(M) and (M, L) are cliques that cover the graph.

We define $\mathcal{Z}$ as the regulating constant in the formula. we have what follows :

$$\mathcal{Z} = \sum_{A,B,D,F,T,L,M,X} \mathcal{F}(A)\mathcal{F}(A,T)\mathcal{F}(T,F,L)\mathcal{F}(F,X)\mathcal{F}(F,D,B)\mathcal{F}(B,M)\mathcal{F}(M)\mathcal{F}(M,L)$$

$$\mathcal{Z} = \sum_{A,B,D,F,T,L,M,X} p(A, B, D, F, T, L, M, X) = 1$$

Thus we have $\mathcal{Z} = 1$

Factor graph

we define the following set of factor functions :

$$\begin{cases} \psi(A) = & p(A) \\ \psi(A,T) = & p(T|A) \\ \psi(T,F,L) = & p(F|T,L) \\ \psi(F,X) = & p(X|F) \\ \psi(F,D,B) = & p(D|F,B) \\ \psi(B,M) = & p(B|M) \\ \psi(M) = & p(M) \\ \psi(M,L) = & p(L|M) \end{cases} \tag{3.2}$$

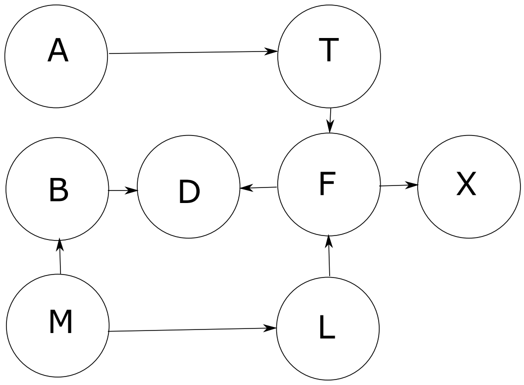

Figure 2 : Directed Graph for given factorization

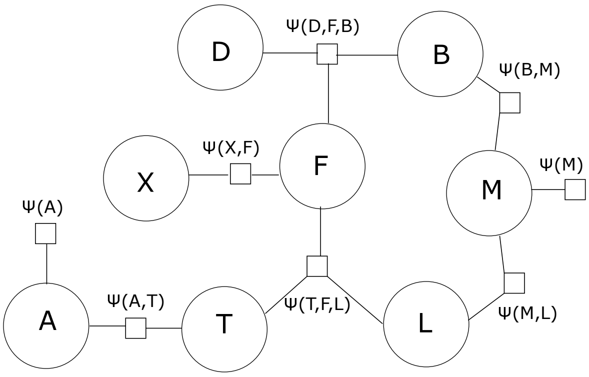

Figure 3 : Factorized graph for given factor function

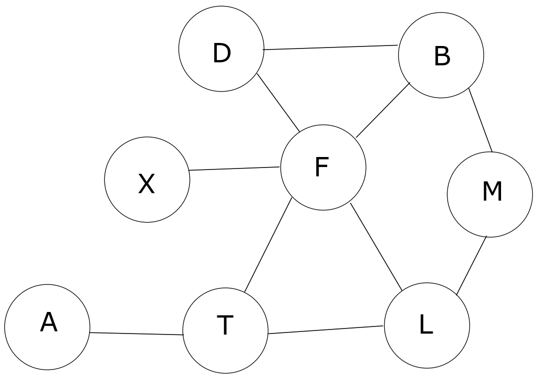

Figure 4 : Undirected graph for given clique function

Memory allocation

Under the given factorization, let $\mathcal{M}$ be the space required. Each variable has N possibility. If we use the non-factorized probability we need $\mathcal{O}(N^8)$ for memory.

$$\mathcal{M} = N^3 + N + N^2 + N^2 + N^2 + N^2 + N^3 + N$$

$$\mathcal{M} = \mathcal{O}(N^3) << \mathcal{O}(N^8)$$

Dependency

$A \perp M|\emptyset$ **?**

$$\nu = \emptyset$$

existing undirected paths :

$$P_1 : A \to T \to F \to L \to M$$

$$P_2 : A \to T \to F \to D \to B \to M$$

P_1 and P_2 are blocked by ν because :

$$\begin{cases} D \in P_1 \\ \quad D \; has \; converging \; arrows \\ \quad D \notin \nu \end{cases} \qquad \begin{cases} F \in P_2 \\ \quad F \; has \; converging \; arrows \\ F \Rightarrow X \\ \quad F \notin \nu \\ \quad X \notin \nu \end{cases} \tag{3.3}$$

$$\nu \quad d - separates \; A \; from \; M$$

$$\boxed{A \perp M|\emptyset}$$

$A \perp M|X$ **?**

$$\nu = \{X\}$$

existing undirected paths :

$$P_1 : A \to T \to F \to L \to M$$

$$P_2 : A \to T \to F \to D \to B \to M$$

P_1 isn't blocked by ν because :
— all nodes with converging arrows are or have descendant in ν
— all nodes without converging arrows aren't elements of ν

$$\nu \quad doesn't \; d - separates \; A \; from \; M$$

$$\boxed{A \not\perp M|X}$$

$T \perp L|X$ **?**

$$\nu = \{X\}$$

existing undirected paths :

$$P_1 : T \rightarrow F \rightarrow L$$
$$P_2 : T \rightarrow F \rightarrow D \rightarrow B \rightarrow M \rightarrow L$$

P_1 isn't blocked by ν because :
— Descendant of F is in ν (observed).

$$\nu \quad doesn't\ d - separates\ T\ from\ L$$
$$\boxed{T \not\perp L|X}$$

$X \perp L|F$ **?**

$$\nu = \{F\}$$

existing undirected paths :

$$P_1 : X \rightarrow F \rightarrow L$$
$$P_2 : X \rightarrow F \rightarrow D \rightarrow B \rightarrow M \rightarrow L$$

P_1 and P_2 are blocked by ν because :

$$\begin{cases} F \in P_1 \\ F\ does\ not\ have\ converging\ arrows \\ F \in \nu \end{cases} \qquad \begin{cases} D \in P_2 \\ D\ has\ converging\ arrows\ and\ no\ desendant \\ D \notin \nu \end{cases} \tag{3.4}$$

$$\nu \quad d - separates\ X\ from\ L$$
$$\boxed{X \perp L|F}$$

$X \perp L|D$ **?**

$$\nu = \{D\}$$

existing undirected paths :

$$P_1 : X \rightarrow F \rightarrow L$$
$$P_2 : X \rightarrow F \rightarrow D \rightarrow B \rightarrow M \rightarrow L$$

P_1 isn't blocked by ν because :
— F does not have converging arrows and is not in ν (not observed).

$$\nu \quad doesn't\ d - separates\ T\ from\ L$$
$$\boxed{X \not\perp L|D}$$

4 Question 4

The Generalized gamma distribution is defined as follow with parameters (α, β, c) :

$$\mathcal{GG}(v; \alpha, \beta, c) = \frac{|c|}{\Gamma(\alpha)\beta^{c\alpha}} v^{c\alpha-1} \exp(-(v/\beta)^c)$$

We want to check if the Generalized Gamma distribution is an exponential family. Let $\theta = (\alpha, \beta, c)$.

We have

$$\mathcal{GG}(v; \alpha, \beta, c) = C(\theta)h(v) \exp(-\beta^{-c}v^c + c\alpha \ln(v))$$

with $C(\theta) = \frac{|c|}{\Gamma(\alpha)\beta^{c\alpha}}$ and $h(v) = 1/v$

To say that Generalized Gamma distribution is an exponential family we need to write $\exp(-\beta^{-c}v^c + c\alpha \ln(v))$ as $\exp(< \eta(\theta), T(v) >)$

Given that we have the term v^c, we cannot separate v and θ in order to get function η and T. In other terms, v^c cannot be written as a scalar product of function of v and θ. Therefore the Generalized Gamma distribution is not an exponential family.

Famous distributions from Generalized Gamma

- Inverse Gamma Distribution
  By taking $c = -1$ we get

$$\mathcal{IG}(v, \alpha, \beta) = \mathcal{GG}(v, \alpha, \beta, 1) = \frac{1}{\Gamma(\alpha)} \frac{v^{-(\alpha+1)}}{\beta^{-\alpha}} \exp(-(\beta/v))$$

- Exponential distribution
  For $c = 1$ and $\alpha = 1$ and knowing that $\Gamma(1) = 1/$ We get the exponential distribution of parameter $\lambda = 1/\beta$

$$\mathcal{E}(v, \lambda) = \lambda \exp(-\lambda v) = \mathcal{GG}(v, 1, 1/\beta, 1)$$