Report assignement 1A - Machine Learning Advanced DD2434

Flandre Corentin

from November 14, 2023 to December 8, 2023

1 1B Assignement

1.1 CAVI for Earth quakes

Question 1.1.1: The DGM of the model described for earthquakes in the subject is like this :

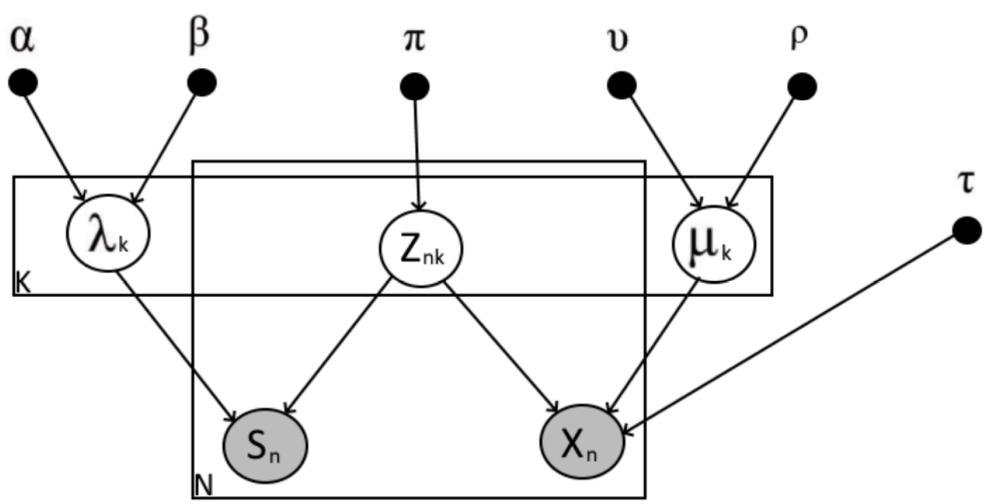

Figure 1: My DGM representation for the earthquakes model

Question 1.1.2: First, the expression $p(X, S, Z, \lambda, \mu | \pi, \tau, \alpha, \beta, v, \rho)$ can be written like this, using conditional probability/the Bayes rules :

$$p(S|Z, \lambda) * p(\lambda|\alpha, \beta) * p(Z|\pi) * p(X|Z, \mu, \tau) * p(\mu, |v, \rho)$$

so with K super-epicentra and N observations :

$$p(X, S, Z, \lambda, \mu|\pi, \tau, \alpha, \beta, v, \rho) \quad = \quad \prod_{k=1}^{K} p(\mu_k, v, \rho) * p(\lambda_k|\alpha, \beta) \prod_{n=1}^{N} p(Z_n = k) * p(S_n|Z = k, \lambda_k) * p(X_n|Z_n = k, \mu_k, \tau_k)$$

with the given and known distribution :

$$\begin{aligned} Z_n &\quad follows \quad Cat(\pi) \\ S_n|Z_n = k, \lambda_k &\quad follows \quad Poisson(\lambda_k) \\ X_n|Z_n = k, \mu_k, \tau &\quad follows \quad N_{2D}(\mu_k, \tau) \\ \mu_k &\quad follows \quad N_{2D}(v, \rho) \\ \lambda_k &\quad follows \quad Gamma(\alpha, \beta) \end{aligned}$$

Secondly, we can use the pdfs/pmfs :

$$\begin{aligned} & p(X, S, Z, \lambda, \mu|\pi, \tau, \alpha, \beta, v, \rho) \\ = & \prod_{k=1}^{K} \frac{\lambda_k^{S_n}}{S_n!} \frac{1}{\sqrt{(2\pi)^2|\tau|}} e^{\frac{-1}{2}(X_n-\mu_k)^T\tau^{-1}(X_n-\mu_k)} \prod_{n=1}^{N} \frac{\lambda_k^{\alpha-1}\beta^{\alpha}e^{-\beta\lambda_k}}{\Gamma(\alpha)} \frac{1}{\sqrt{(2\pi)^2|\rho|}} e^{\frac{-1}{2}(\mu_k-v)^T\rho^{-1}(\mu_k-v)}\pi_{nk} \end{aligned}$$

Then, we want the log expression :

$$\begin{aligned} & \log p(X, S, Z, \lambda, \mu|\pi, \tau, \alpha, \beta, v, \rho) \\ = & \sum_{k=1}^{K}[S_n \log \lambda_k - \log S_n! - \log 2 - \frac{1}{2}\log|\tau| - \frac{1}{2}(X_n - \mu_k)^T\tau^{-1}(X_n - \mu_k) + \sum_{n=1}^{N}[(\alpha-1)\log\lambda_k + \alpha\log\beta - \beta\lambda_k \\ & - \log\Gamma(\alpha) - \log 2\pi - \frac{1}{2}\log|\rho| - \frac{1}{2}(\mu_k - v)^T\rho^{-1}(\mu_k - v) + \log\pi_{nk}]] \end{aligned}$$

Since, $\pi, \tau, \alpha, \beta, v, \rho$ are treated as constants (subject) :

$$\boxed{\begin{aligned} & \log p(X, S, Z, \lambda, \mu|\pi, \tau, \alpha, \beta, v, \rho) \\ = & \sum_{k=1}^{K}[S_n \log \lambda_k - \log S_n! - \frac{1}{2}(X_n - \mu_k)^T\tau^{-1}(X_n - \mu_k) + \sum_{n=1}^{N}[(\alpha-1)\log\lambda_k - \beta\lambda_k - \frac{1}{2}(\mu_k - v)^T\rho^{-1}(\mu_k - v)]] \end{aligned}}$$

Question 1.1.3: Firstly, with the assumption $q(Z, \mu, \lambda) = \prod_{n=1}^{N} q(Z_n) \prod_{k=1}^{K} q(\mu_k)q(\lambda_k)$, we can compute the CAVI update equation step by step :

$$\begin{aligned} q^*(Z_n) &= expE_{-Z_n}[(X_n, S_n, Z_n = k, \mu_k, \lambda_k)] \\ => log(q^*(Z_n)) &= E_{-Z_n}[\log p(S_n|Z_n = k, \lambda_k) + \log p(\lambda_k|\alpha, \beta) + \log p(Z_n = k|\pi) + \log p(X_n|Z_n = k, \mu_k, \tau) + \log p(\mu_k, |v, \rho)] \\ &= E_{-Z_n}[\log p(S_n|Z_n = k, \lambda_k) + \log p(Z_n = k|\pi) + \log p(X_n|Z_n = k, \mu_k, \tau)] + cste1 \\ &= E_{\lambda_k,\mu_k}[\log p(S_n|Z_n = k, \lambda_k) + \log p(X_n|Z_n = k, \mu_k, \tau)] + \log p(Z_n = k|\pi) + cste1 \\ &= E_{\lambda_k,\mu_k}[S_n \log\lambda_k - \lambda_k - \log S_n! - \frac{1}{2}(X_n - \mu_k)^T\tau^{-1}(X_n - \mu_k)] + \log\pi_{n,k} + cste2 \\ &= S_n E_{\lambda_k}[\log\lambda_k] - E_{\lambda_k}[\lambda_k] - \log S_n! - \frac{1}{2}(X_n - E_{\mu_k}[\mu_k])^T\tau^{-1}(X_n - E_{\mu_k}[\mu_k])] + \log\pi_{n,k} + cste2 \end{aligned}$$

$$
\begin{aligned}
q^*(\mu_k) &= expE_{-\mu_k}[(X_n, S_n, Z_n = k, \mu_k, \lambda_k)] \\
=> log(q^*(\mu_k)) &= E_{-\mu_k}[\log p(S_n|Z_n = k, \lambda_k) + \log p(\lambda_k|\alpha, \beta) + \log p(Z_n = k|\pi) + \log p(X_n|Z_n = k, \mu_k, \tau) + \log p(\mu_k, |v, \rho)] \\
&= E_{-\mu_k}[\log p(X_n|Z_n = k, \mu_k, \tau) + \log p(\mu_k, |v, \rho)] + cste1 \\
&= E_{Z_n}[\log p(X_n|Z_n = k, \mu_k, \tau)] + \log p(\mu_k, |v, \rho) + cste1 \\
&= E_{Z_n}[-\frac{1}{2}(X_n - \mu_k)^T \tau^{-1}(X_n - \mu_k)] + -\frac{1}{2}(\mu_k - v)^T \rho^{-1}(\mu_k - v) + cste2 \\
&= -\frac{1}{2}(X_n - \mu_k)^T \tau^{-1}(X_n - \mu_k) + -\frac{1}{2}(\mu_k - v)^T \rho^{-1}(\mu_k - v) + cste2
\end{aligned}
$$

$$
\begin{aligned}
q^*(\lambda_k) &= expE_{-\lambda_k}[(X_n, S_n, Z_n = k, \mu_k, \lambda_k)] \\
=> log(q^*(\lambda_k)) &= E_{-\lambda_k}[\log p(S_n|Z_n = k, \lambda_k) + \log p(\lambda_k|\alpha, \beta) + \log p(Z_n = k|\pi) + \log p(X_n|Z_n = k, \mu_k, \tau) + \log p(\mu_k, |v, \rho)] \\
&= E_{-\lambda_k}[\log p(S_n|Z_n = k, \lambda_k) + \log p(\lambda_k|\alpha, \beta)] + cste1 \\
&= E_{Z_n}[\log p(S_n|Z_n = k, \lambda_k)] + \log p(\lambda_k|\alpha, \beta) + cste1 \\
&= E_{Z_n}[S_n \log \lambda_k - \lambda_k - \log S_n!] + (\alpha - 1) \log \lambda_k - \beta\lambda_k + cste2 \\
&= S_n \log \lambda_k - \lambda_k - \log S_n! + (\alpha - 1) \log \lambda_k - \beta\lambda_k + cste2
\end{aligned}
$$

Now we have expression of update assumption for each variable in function of observations, variable previous iterations and parameters treated as constants

1.2 VAE image generation

My results for the VAE image generation are not what is expected by the subject, I prefer not to give them.