CS5340 - Lab2 Part2 Report

Liu Yichao A0304386A

September 23, 2024

Question We suppose that the full observation is generated by a linear-gaussian model.

$$p(x_u|x_{\pi_u}, \theta_u) = \mathcal{N}_u[w_{u0}, \dots, w_{uC}, \sigma_u^2] = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left\{-0.5\frac{\left(x_u - \left(\sum_{c \in x_{\pi_u}} w_{uc}x_{uc} + w_{u0}\right)\right)^2}{\sigma_u^2}\right\} \tag{1}$$

To derive the MLE estimates of θ, we can equivalently find the maximum log-likelihood estimate:

$$\arg\max_{\theta_u} \sum_{n=1}^{N} \log p(x_u|x_{\pi_u}, \theta_u|x_{\pi_u}) =$$

$$\arg\max_{\theta_u} \underbrace{\sum_{n=1}^{N} \left\{-\frac{1}{2}\log 2\pi\sigma_u^2 - \frac{1}{2\sigma_u^2}\left(x_{u,n} - \left(\sum_{c \in x_{\pi_u}} w_{uc}x_{uc,n} + w_{u0}\right)\right)^2\right\}}_{\mathcal{L}} \tag{2}$$

Compute the MLE estimates of $\theta_u = (W, \sigma_u^2)$.

Derivation of W Let $W = \begin{bmatrix} w_{u0} \\ w_{u1} \\ \vdots \\ w_{uC} \end{bmatrix}$, Take the derivative of $\mathcal{L}$ with respect to w_{u0}:

$$\frac{\partial \mathcal{L}}{\partial w_{u0}} = \sum_{n=1}^{N} \left\{\frac{1}{\sigma_u^2}\left(x_{u,n} - \left(\sum_{c \in x_{\pi_u}} w_{uc}x_{uc,n} + w_{u0}\right)\right)\right\} = 0 \tag{3}$$

Take the derivative of $\mathcal{L}$ with respect to $w_{uc}, c = 1, \dots, C$, C is the number of parents of x_u:

$$\frac{\partial \mathcal{L}}{\partial w_{uc}} = \sum_{n=1}^{N} \left\{\frac{1}{\sigma_u^2}\left(x_{u,n} - \left(\sum_{c \in x_{\pi_u}} w_{uc}x_{uc,n} + w_{u0}\right)\right)x_{uc,n}\right\} = 0 \tag{4}$$

Let $Y = \begin{bmatrix} \sum_{n=1}^N x_{u,n} \\ \sum_{n=1}^N x_{u,n}x_{u1,n} \\ \vdots \\ \sum_{n=1}^N x_{u,n}x_{uC,n} \end{bmatrix}$, $A = \begin{bmatrix} 1 & x_{u1,1} & x_{u2,1} & \cdots & x_{uC,1} \\ 1 & x_{u1,2} & x_{u2,2} & \cdots & x_{uC,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{u1,N} & x_{u2,N} & \cdots & x_{uC,N} \end{bmatrix}$,

$X = A^T A = \begin{bmatrix} N & \sum_{n=1}^N x_{u1,n} & \cdots & \sum_{n=1}^N x_{uC,n} \\ \sum_{n=1}^N x_{u1,n} & \sum_{n=1}^N x_{u1,n}^2 & \cdots & \sum_{n=1}^N x_{u1,n}x_{uC,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{n=1}^N x_{uC,n} & \sum_{n=1}^N x_{uC,n}x_{u1,n} & \cdots & \sum_{n=1}^N x_{uC,n}^2 \end{bmatrix}$,

then we can rewrite eq. (3) and eq. (4) as

$$XW = Y$$

Since each line of A is an observation of the parents of x_u, we can assume A is full rank. Thus $X = A^T A$ is also full rank. So the solution is

$$W = X^{-1} Y$$

Derivation of σ_u^2 Take the derivative of $\mathcal{L}$ to σ_u^2:

$$\frac{\partial \mathcal{L}}{\partial \sigma_u^2} = \sum_{n=1}^{N} \left\{ -\frac{1}{2\sigma_u^2} + \frac{1}{2\sigma_u^4} \left(x_{u,n} - \left(\sum_{c \in x_{\pi_u}} w_{uc} x_{uc,n} + w_{u0} \right) \right)^2 \right\} = 0 \tag{5}$$

Let $Z = \sum_{n=1}^{N} \left(x_{u,n} - \left(\sum_{c \in x_{\pi_u}} w_{uc} x_{uc,n} + w_{u0} \right) \right)^2$, then we can rewrite eq. (5) as

$$\frac{N}{2\sigma_u^2} = \frac{Z}{2\sigma_u^4}$$

so the solution is

$$\sigma_u^2 = \frac{Z}{N}$$

Implementation The code is just an implementation of the above derivation.