

# Problem class sheet 4 (DRAFT, TO BE REFINED AFTER THE PROBLEM CLASS)

Lecturer &amp; author: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

**Exercise 1.** (★★) Consider the model

$$Z = BZ + (I - B)X\beta + E$$

where  $E \sim N(0, \sigma^2 I)$ ,  $X$  is a  $n \times p$  design matrix  $X$ ,  $\beta \in \mathbb{R}^p$ ,  $B$  is an  $n \times n$  matrix with  $[B]_{i,i} = 0$ , and  $(I - B)$  is non-singular.

(1) Show that

$$E(Z) = X\beta$$

$$\text{Var}(Z) = \sigma^2 (I - B)^{-1} (I - B^\top)^{-1}$$

(2) Show that the above model is SAR for  $Z - E(Z)$

(3) Compute the Maximum Likelihood Estimators (MLE)  $\hat{\beta}$ , and  $\hat{\sigma}^2$  of  $\beta$ , and  $\sigma^2$ . Assume that  $((I - B)X)^\top (I - B)X$  is non-singular.

(4) Find the sampling distribution of  $\hat{\beta}$  given  $X$ .

**Solution.**

(1) It is

$$E(Z) = E(BZ + (I - B)X\beta + E) \iff$$

$$E(Z) = E(BZ) + (I - B)X\beta + E(E) \iff$$

$$(I - B)E(Z) = (I - B)X\beta + \overset{=0}{\cancel{E(E)}} \iff$$

$$E(Z) = X\beta$$

also

$$\text{Var}((I - B)Z) = \text{Var}((I - B)X\beta + E)$$

$$\text{Var}((I - B)Z) = \text{Var}(E)$$

$$(I - B)\text{Var}(Z)(I - B)^\top = \sigma^2 I$$

$$\text{Var}(Z) = (I - B)^{-1} \sigma^2 I (I - B)^{-\top}$$

(2) It is

$$Z - E(Z) = B(Z - X\beta) + E$$

where  $E \sim N(0, \sigma^2 I)$ , hence  $Z - E(Z)$  is a SAR model given the assumptions taken.  
(3) The likelihood of  $Z$  given the parameters  $\beta$ , and  $\sigma^2$  is

$$\begin{aligned} L(Z; \beta, \sigma^2) &= N(Z|E(Z), \text{Var}(Z)) \\ &= N\left(Z|X\beta, (I - B)^{-1} \sigma^2 I (I - B)^{-\top}\right) \end{aligned}$$

Hence

$$\begin{aligned} -2 \log(L(Z; \beta, \sigma^2)) &= -2 \log\left(N\left(Z|X\beta, (I - B)^{-1} \sigma^2 I (I - B)^{-\top}\right)\right) \\ &= \log\left(\det\left((I - B)^{-1} \sigma^2 I (I - B)^{-\top}\right)\right) \\ &\quad + (Z - X\beta)^\top \left((I - B)^{-1} \sigma^2 I (I - B)^{-\top}\right)^{-1} (Z - X\beta) \\ &= \log\left(\det\left((I - B)^{-1} \sigma^2 I (I - B)^{-\top}\right)\right) \\ &\quad + \frac{1}{\sigma^2} (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta) \end{aligned}$$

The likelihood equations are

$$\begin{aligned} 0 &= \nabla_{(\beta, \sigma^2)} (-2 \log(L(Z; \beta, \sigma^2))) \Big|_{(\beta, \sigma^2) = (\hat{\beta}, \hat{\sigma}^2)} \\ &= \left[ \begin{array}{c} \frac{\partial}{\partial \beta} (-2 \log(L(Z; \beta, \sigma^2))) \\ \frac{\partial}{\partial \sigma^2} (-2 \log(L(Z; \beta, \sigma^2))) \end{array} \right]_{(\beta, \sigma^2) = (\hat{\beta}, \hat{\sigma}^2)} \\ &= \left[ \begin{array}{c} -X^\top 2 (I - B)^\top (I - B) (Z - X\beta) \\ -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta) \end{array} \right]_{(\beta, \sigma^2) = (\hat{\beta}, \hat{\sigma}^2)} \end{aligned}$$

This is because

$$\begin{aligned} \frac{\partial}{\partial \beta} (-2 \log(L(Z; \beta, \sigma^2))) &= \frac{\partial}{\partial \beta} \left( \frac{1}{\sigma^2} (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta) \right) \\ &= \left[ \frac{\partial}{\partial \beta} (Z - X\beta) \right] \frac{\partial}{\partial \xi} \left( \frac{1}{\sigma^2} (\xi)^\top (I - B)^\top (I - B) (\xi) \right) \Big|_{\xi = Z - X\beta} \\ &= -X^\top 2 (I - B)^\top (I - B) (Z - X\beta) \end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial \sigma^2} (-2 \log (L(Z; \beta, \sigma^2))) &= \frac{\partial}{\partial \sigma^2} \left( \log \left( \det \left( (I - B)^{-1} \sigma^2 I (I - B)^{-\top} \right) \right) \right) \\
&\quad + \frac{\partial}{\partial \sigma^2} \left( \frac{1}{\sigma^2} (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta) \right) \\
&= \frac{\partial}{\partial \sigma^2} (-n \log (\sigma^2)) + \frac{\partial}{\partial \sigma^2} \left( \frac{1}{\sigma^2} \right) (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta) \\
&= -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta)
\end{aligned}$$

So the likelihood equations are

$$\begin{aligned}
0 &= X^\top (I - B)^\top (I - B) (Z - X\hat{\beta}) \\
0 &= -\frac{n}{\hat{\sigma}^2} + \frac{1}{\hat{\sigma}^4} (Z - X\hat{\beta})^\top (I - B)^\top (I - B) (Z - X\hat{\beta})
\end{aligned}$$

Solving the first equation wrt  $\hat{\beta}$  I get

$$\begin{aligned}
0 &= X^\top (I - B)^\top (I - B) (Z - X\hat{\beta}) \iff \\
0 &= X^\top (I - B)^\top (I - B) Z - X^\top (I - B)^\top (I - B) X\hat{\beta} \iff \\
X^\top (I - B)^\top (I - B) X\hat{\beta} &= X^\top (I - B)^\top (I - B) Z \iff \\
\hat{\beta} &= \left( X^\top (I - B)^\top (I - B) X \right)^{-1} X^\top (I - B)^\top (I - B) Z
\end{aligned}$$

provided that  $X^\top (I - B)^\top (I - B) X$  is non-singular (this is given, anyway).

Solving the second equation wrt  $\hat{\sigma}^2$  I get

$$\begin{aligned}
0 &= -\frac{n}{\hat{\sigma}^2} + \frac{1}{\hat{\sigma}^4} (Z - X\hat{\beta})^\top (I - B)^\top (I - B) (Z - X\hat{\beta}) \iff \\
0 &= -\frac{n}{1} + \frac{1}{\hat{\sigma}^2} (Z - X\hat{\beta})^\top (I - B)^\top (I - B) (Z - X\hat{\beta}) \iff \\
\hat{\sigma}^2 &= \frac{1}{n} (Z - X\hat{\beta})^\top (I - B)^\top (I - B) (Z - X\hat{\beta})
\end{aligned}$$

- (4) It is Normal as a linear combination of Normally distributed random variables. Its moments (mean and variance) are

$$\begin{aligned}
E(\hat{\beta}|X) &= E \left( \left( X^\top (I - B)^\top (I - B) X \right)^{-1} X^\top (I - B)^\top (I - B) Z | X \right) \\
&= \left( X^\top (I - B)^\top (I - B) X \right)^{-1} X^\top (I - B)^\top (I - B) E(Z|X) \\
&= \left( X^\top (I - B)^\top (I - B) X \right)^{-1} X^\top (I - B)^\top (I - B) X\beta \\
&= \beta
\end{aligned}$$

and

$$\begin{aligned}
\text{Var}(\hat{\beta}|X) &= \text{Var}\left(\left(X^\top (I - B)^\top (I - B) X\right)^{-1} X^\top (I - B)^\top (I - B) Z|X\right) \\
&= \left(X^\top (I - B)^\top (I - B) X\right)^{-1} X^\top (I - B)^\top (I - B) \text{Var}(Z|X) \\
&\quad \left(\left(X^\top (I - B)^\top (I - B) X\right)^{-1} X^\top (I - B)^\top (I - B)\right)^\top \\
&= \left(X^\top (I - B)^\top (I - B) X\right)^{-1} X^\top \cancel{(I - B)^\top (I - B)} \\
&\quad \cancel{\sigma^2(I - B)^{-1} (I - B^\top)^{-1}} \\
&\quad \left(\left(X^\top (I - B)^\top (I - B) X\right)^{-1} X^\top (I - B)^\top (I - B)\right)^\top \\
&= \sigma^2 \left(\cancel{X^\top (I - B)^\top (I - B) X}\right)^{-1} X^\top \cancel{(I - B)^\top (I - B) X} \left(X^\top (I - B)^\top (I - B) X\right)^{-1} \\
&= \sigma^2 \left(X^\top (I - B)^\top (I - B) X\right)^{-1}
\end{aligned}$$

- Notice that, in Frequentist Statistical framework, once we have computed the sampling distributions (those above), we can produce inference tools in the similar manner to Normal Linear regression.

**Exercise 2.** (★) Suppose that  $\mathcal{S}$  is a finite set that contains at least two elements and is equipped with a symmetric relation  $\sim$ . Consider the Poisson auto-regression model defined as

$$\begin{cases} y_i | y_{\mathcal{S} \setminus \{i\}} \sim \text{Poisson}(\mu_i) \\ \log(\mu_i) = \theta \sum_{i \sim j, j \neq i} y_j \end{cases}$$

for  $y \in \mathbb{N}^{\mathcal{S}}$ .

**Hint:** You can use that if  $X \sim \text{Poisson}(\mu)$  then  $X$  has PMF

$$\Pr_X(x|\mu) = \frac{1}{x!} \exp(-\mu) \mu^x \mathbf{1}(x \in \{0, 1, 2, \dots\})$$

- (1) Show that the above model is well-defined if and only if  $\theta \leq 0$ .
- (2) Find the canonical potential with respect to  $\zeta = 0$ .

**Solution.** It is

$$\Pr_i(y_i | y_{\mathcal{S} \setminus \{i\}}) = \frac{1}{y_i!} \exp(-\mu_i) \mu_i^{y_i} \mathbf{1}(y_i \in \mathbb{N})$$

- (1) It is

$$\Pr_i(y_i = 0 | y_{\mathcal{S} \setminus \{i\}}) = \exp(-\mu_i) =$$

and

$$\Pr_i(y_i = \ell | y_{S \setminus \{i\}}) = \frac{1}{\ell!} \exp(-\mu_i) \mu_i^\ell$$

for  $\ell \in \mathbb{N}$  Then by the factorization theorem wrt reference 0 it is

$$\begin{aligned} \frac{\Pr_Y(y)}{\Pr_Y(0)} &= \prod_{i \in S} \frac{\Pr_i(y_i | y_1, \dots, y_{i-1}, 0, \dots, 0)}{\Pr_i(0 | y_1, \dots, y_{i-1}, 0, \dots, 0)} \\ &= \prod_{i \in S} \frac{\frac{1}{y_i!} \exp(-\mu_i) \mu_i^{y_i}}{\exp(-\mu_i)} = \prod_{i \in S} \frac{1}{y_i!} \mu_i^{y_i} \\ &= \prod_{i \in S} \frac{1}{y_i!} \exp\left(\theta \sum_{i \sim j, j \neq i} y_j\right)^{y_i} \\ &= \exp\left(\theta \sum_{i \sim j, j \neq i} y_i y_j - \sum_{i \in S} \log(y_i!)\right) \end{aligned}$$

That is

$$\Pr_Y(y) = \exp\left(\theta \sum_{i \sim j, j \neq i} y_i y_j - \sum_{i \in S} \log(y_i!)\right) \Pr_Y(0)$$

Now, if  $\theta \leq 0$  then  $\theta \sum_{i \sim j, j \neq i} y_i y_j \leq 0$  hence the constant is

$$\begin{aligned} \sum_{y \in \mathbb{N}^S} \frac{\Pr_Y(y)}{\Pr_Y(0)} &= \sum_{y \in \mathbb{N}^S} \exp\left(\theta \sum_{i \sim j, j \neq i} y_i y_j - \sum_{i \in S} \log(y_i!)\right) \\ &\leq \sum_{y \in \mathbb{N}^S} \exp\left(-\sum_{i \in S} \log(y_i!)\right) \\ &= \sum_{y \in \mathbb{N}^S} \prod_{i \in S} \frac{1}{y_i!} \\ &= \left(\sum_{\ell=0}^{\infty} \frac{1}{\ell!}\right)^{\text{Card}(S)} < \infty \end{aligned}$$

If  $\theta > 0$  without loss of generality consider the first two sites and suppose that  $1 \sim 2$ , then

$$\frac{\Pr_Y((y_1, y_2, 0, \dots, 0)^\top)}{\Pr_Y(0)} = \frac{\exp(\theta y_1 y_2)}{y_1! y_2!}$$

should be summable. However, the series

$$\sum_{y_1=1}^{\infty} \sum_{y_2=1}^{\infty} \frac{\exp(\theta y_1 y_2)}{y_1! y_2!} = \infty$$

diverges as the general term does not go to zero.

(2) By definition,  $V_\emptyset = 0$ .

Then I will use Theorem ??, and  $\zeta = 0$ .

For  $\mathcal{A} = \{i\}$ , it is

$$V_{\{i\}}(y) = \log \left( \Pr_i(y_i|0, \dots, 0) \right) - \log \left( \Pr_i(0|0, \dots, 0) \right) = -\log(y_i!)$$

For  $\mathcal{A} = \{i, j\}$ , it is

$$\begin{aligned} V_{\{i,j\}}(y) &= \log \left( \Pr_i(y_i|y_j, 0, \dots, 0) \right) \\ &\quad - \log \left( \Pr_i(y_i|0, \dots, 0) \right) - \log \left( \Pr_i(y_j|0, \dots, 0) \right) \\ &\quad + \log \left( \Pr_i(0|0, \dots, 0) \right) \\ &= -y_i y_j 1(i \sim j) \end{aligned}$$

So

$$V_{\{i,j\}}(y) = -y_i y_j 1(i \sim j)$$

Since the joint distribution is proportional such as

$$\begin{aligned} \Pr_Y(y) &= \exp \left( \theta \sum_{i \sim j, j \neq i} y_i y_j - \sum_{i \in \mathcal{S}} \log(y_i!) \right) \Pr_Y(0) \\ &\propto \exp \left( \theta \sum_{i \sim j, j \neq i} y_i y_j - \sum_{i \in \mathcal{S}} \log(y_i!) \right) \end{aligned}$$

all the other potentials are zero.

- Perhaps not the most elegant derivation. In the next lecture, we will use a more general tool to compute such stuff which is based on the “exponential distribution family”.

---

(★) Show

that the local characteristics

$$\Pr_1(x|y) = \Pr_2(x|y) = \sqrt{\frac{1}{2\pi}} \exp \left( -\frac{1}{2} (y - x)^2 \right)$$

do not define a proper joint distribution on  $\mathbb{R}^{\{1,2\}}$

**Solution.**

---