

Homework 4: Aerial data modeling

Lecturer: Georgios P. Karagiannis

georgios.karagiannis@durham.ac.uk

Exercise 1. (★) Show that the local characteristics

$$\begin{aligned}\text{pr}_1(x_1|x_2) &= \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}(x_1 - x_2)^2\right) \\ \text{pr}_2(x_2|x_1) &= \sqrt{\frac{1}{2\pi}} \exp\left(-\frac{1}{2}(x_2 - x_1)^2\right)\end{aligned}$$

do not define a proper joint distribution on $\mathbb{R}^{\{1,2\}}$.**Solution.** The local characteristics are well-defined normal densities. Using Besag's factorization theorem, with reference $x^* = (0, 0)^\top$, the joint density would be proportional to

$$\frac{\text{pr}_1(x_1|0)}{\text{pr}_1(0|0)} \frac{\text{pr}_2(x_2|x_1)}{\text{pr}_2(0|x_1)} = \exp\left(-\frac{1}{2}(x_2 - x_1)^2\right) = G(x_1, x_2)$$

for $(x, y) \in \mathbb{R}^2$. But $G(x, y)$ is not integratable,

$$\int_{\mathbb{R}^2} G(x_1, x_2) dx_1 dx_2 = \int_{\mathbb{R}^2} \exp\left(-\frac{1}{2}(x_2 - x_1)^2\right) dx_1 dx_2 = \infty$$

Definition. (Pseudo-likelihood) The pseudo likelihood pseudo- $L(Z; \theta)$ of observables $Z = (Z_1, \dots, Z_n)^\top$ given parameters θ is an approximation of the likelihood $L(Z; \theta)$ of observables $Z = (Z_1, \dots, Z_n)^\top$ given parameters θ which is equal to

$$\text{pseudo-}L(Z; \theta) = \prod_i \text{pr}(Z_i | Z_{-i}, \theta)$$

where $\text{pr}(Z_i | Z_{-i}, \theta)$ are the conditionals of the joint pdf/pmf of the sampling distribution $\text{pr}(Z | \theta)$ of Z given parameter θ .**Definition.** (Maximum Pseudo-Likelihood Estimator) The Maximum Pseudo-Likelihood Estimator (MPLE) $\tilde{\theta}$ of θ is the maximizer of the pseudo likelihood function pseudo- $L(Z; \theta)$ where the parameter θ is the argument and the observables $Z = (Z_1, \dots, Z_n)^\top$ are fixed values.

$$\tilde{\theta} = \arg \max_{\theta} (\text{pseudo-}L(Z; \theta))$$

Exercise 2. (★) Let $Z \in \mathcal{Z}^{\mathcal{S}}$ where $\mathcal{S} = \{1, \dots, n\}$ and $\mathcal{Z} = \mathbb{R}$. Consider the model

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

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

Hint: The following formulas are provided for your information

- $\partial (XY) = (\partial X) Y + X (\partial Y)$
- $\partial (X^\top) = (\partial X)^\top$
- $\frac{\partial}{\partial x} (x^\top B x) = (B + B^\top) x$
- $\frac{\partial}{\partial x} \left((s - Ax)^\top W (s - Ax) \right) = -2AW (s - Ax)$

- (1) This is a multiple choice question, choose any number of correct answers.
 - (a) Z follows a simultaneous autoregressive (SAR) with Gaussian joint distribution with mean $X\beta$ and covariance matrix $\sigma^2 (I - B)^{-1}$
 - (b) Ising model
 - (c) Conditional autoregressive (CAR) with Gaussian joint distribution with mean $X\beta$ and covariance matrix $\sigma^2 I$
 - (d) Bernoulli regression
- (2) Show that the minus two log Pseudo-Likelihood is such as

$$-2 \log (\text{pseudo-}L(Z; \beta, \sigma^2)) = n \log (\sigma^2) + \frac{1}{\sigma^2} (Z - X\beta)^\top (I - B)^2 (Z - X\beta) + \text{const.}$$

- (3) Compute the Maximum Pseudo-Likelihood Estimators (MPLE) $\tilde{\beta}$ and $\tilde{\sigma}^2$ of β and σ^2

Solution.

- (1) The correct answer is (a)
- (2) It is

$$Z | \beta, \sigma^2 \sim N \left(X\beta, (I - B)^{-1} \sigma^2 \right)$$

which is in a CAR model form with $K = \text{diag}(\sigma^2, \dots, \sigma^2)$, and $\mu = X\beta$. Hence the local characteristics are

$$Z_i | Z_{-i}, \beta, \sigma^2 \sim N \left([X\beta]_i + \sum_{i \neq j} B_{i,j} (Z_j - [X\beta]_j), \sigma^2 \right)$$

The -log pseudo-likelihood is

$$\begin{aligned}
-2 \log \left(\prod_{i \in \mathcal{S}} \text{pr}_i (Z_i | Z_{-i}, \beta, \sigma^2) \right) &= -2 \sum_{i \in \mathcal{S}} \log (\text{pr}_i (Z_i | Z_{-i}, \beta, \sigma^2)) \\
&= n \log (\sigma^2) + \sum_{i \in \mathcal{S}} \frac{1}{\sigma^2} \left(Z_i - [X\beta]_i - \sum_{i \neq j} B_{i,j} (Z_j - [X\beta]_j) \right)^2 + \text{const.} \\
&\quad \{\text{let } A := I - B\} \\
&= n \log (\sigma^2) + \sum_{i \in \mathcal{S}} \frac{1}{\sigma^2} \left(\sum_i A_{i,j} (Z_j - [X\beta]_j) \right)^2 + \text{const.} \\
&= n \log (\sigma^2) + \frac{1}{\sigma^2} \sum_{i \in \mathcal{S}} (A_{i,\text{all}} (Z - X\beta))^2 + \text{const.} \\
&= n \log (\sigma^2) + \frac{1}{\sigma^2} [A (Z - X\beta)]^\top [A (Z - X\beta)] + \text{const.} \\
&= n \log (\sigma^2) + \frac{1}{\sigma^2} (Z - X\beta)^\top (I - B)^\top (I - B) (Z - X\beta) + \text{const.} \\
&= n \log (\sigma^2) + \frac{1}{\sigma^2} (Z - X\beta)^\top (I - B)^2 (Z - X\beta) + \text{const.}
\end{aligned}$$

where $(I - B)^2 = (I - B)^\top (I - B)$.

(3) The psudo-likelihood equations are

$$\begin{aligned}
0 &= \nabla_{(\beta, \sigma^2)} \left(-2 \sum_{i \in \mathcal{S}} \log (\text{pr}_i (Z_i | Z_{-i}, \beta, \sigma^2)) \right) \Big|_{(\beta, \sigma^2) = (\tilde{\beta}, \tilde{\sigma}^2)} \\
&= \left[\begin{array}{c} \frac{\partial}{\partial \beta} (-2 \sum_{i \in \mathcal{S}} \log (\text{pr}_i (Z_i | Z_{-i}, \beta, \sigma^2))) \\ \frac{\partial}{\partial \sigma^2} (-2 \sum_{i \in \mathcal{S}} \log (\text{pr}_i (Z_i | Z_{-i}, \beta, \sigma^2))) \end{array} \right]_{(\beta, \sigma^2) = (\tilde{\beta}, \tilde{\sigma}^2)} \\
&= \left[\begin{array}{c} -\frac{1}{\sigma^2} X^\top 2 (I - B)^2 (Z - X\beta) \\ -\frac{n}{\sigma^2} + \frac{1}{\sigma^4} (Z - X\beta)^\top (I - B)^2 (Z - X\beta) \end{array} \right]_{(\beta, \sigma^2) = (\tilde{\beta}, \tilde{\sigma}^2)}
\end{aligned}$$

So the likelihood equations are

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

In details, by solving the first equation wrt $\tilde{\beta}$

$$\begin{aligned}
0 &= \frac{1}{\tilde{\sigma}^2} X^\top (I - B)^2 (Z - X\tilde{\beta}) \Leftrightarrow \\
X^\top (I - B)^2 X \tilde{\beta} &= X^\top (I - B)^2 Z \Leftrightarrow \\
\tilde{\beta} &= \left(X^\top (I - B)^2 X \right)^{-1} X^\top (I - B)^2 Z
\end{aligned}$$

and by solving the second equation wrt $\tilde{\sigma}^2$

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

Hence by solving w.r.t. $\tilde{\beta}$ and $\tilde{\sigma}^2$ I get

$$\begin{aligned}
\tilde{\beta} &= \left(X^\top (I - B)^2 X \right)^{-1} X^\top (I - B)^2 Z \\
\tilde{\sigma}^2 &= \frac{1}{n} \left(Z - X\tilde{\beta} \right)^\top (I - B)^2 \left(Z - X\tilde{\beta} \right)
\end{aligned}$$
