Homework 4: Aerial data modeling

Lecturer: Georgios P. Karagiannis georgios.karagiannis@durham.ac.uk

---

Exercise 1. ($\star$) Show that the local characteristics

$$\mathrm{pr}_1\left(x_1|x_2\right) = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left(x_1-x_2\right)^2\right)$$
$$\mathrm{pr}_2\left(x_2|x_1\right) = \sqrt{\frac{1}{2\pi}}\exp\left(-\frac{1}{2}\left(x_2-x_1\right)^2\right)$$

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

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

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

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

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

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

where $\mathrm{pr}(Z_i|Z_{-i},\theta)$ are the conditionals of the joint pdf/pmf of the sampling distribution $\mathrm{pr}(Z|\theta)$ 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 $Z = (Z_1, ..., Z_n)^\top$ are fixed values.

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

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

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

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

Hint: The following formulas are provided for your information

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

(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

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

(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

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

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

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

The -log pseudo-likelihood is

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

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

(3) The psudo-likelihood equations are

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

So the likelihood equations are

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

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

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

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

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