Handling spatial dependence under unknown unit locations

Supplementary material

1 Proof of Equation (7)

Proof. According to LeSage and Pace (2009, p. 47), the log-likelihood function of the SLM (2) is:

\[
\ln L(\rho; W) = -\frac{n}{2} \ln (2\pi \sigma^2) + \ln \det(A) - \frac{1}{2\sigma^2} (Ay - X\beta)^T(Ay - X\beta), \tag{SM.1}
\]

where \(A = I_n - \rho W\), as defined in the paper.

If (SM.1) is concentrated with respect to \(\rho\), we have (cf. LeSage and Pace 2009, p. 48):

\[
\ln L_c(\rho; W) = -\frac{n}{2} \ln \frac{2\pi e}{\sigma^2} + \ln \det(A) - \frac{n}{2} \ln [(MAy)^T(MAy)], \tag{SM.2}
\]

where \(M = I_n - X(X^TX)^{-1}X^T\) is the projection matrix defined in the paper.

Analogously, the concentrated log-likelihood of model (5) is

\[
\ln L_c(\rho_X; W_X) = -\frac{n}{2} \ln \frac{2\pi e}{\sigma^2} + \ln \det(A_X) - \frac{n}{2} \ln [(MA_Xy)^T(MA_Xy)]. \tag{SM.3}
\]

Equation (7) is obtained by subtracting Equation (SM.2) from (SM.3).

2 Proof of Equation (8)

Proof. In case of squared real matrices (as spatial weight matrices are), the Frobenius norm \(\|\cdot\|_F\) is defined as \(\|A\|_F = \sqrt{\text{tr}(AA^T)}\), for some \(A \in \mathbb{R}^{n \times n}\) (Horn and Johnson 2013, p. 341). Thus, properties of the trace of matrices imply that:

\[
\|\rho_X W_X - \rho W\|_F^2 = \text{tr}((\rho_X W_X - \rho W)(\rho_X W_X - \rho W)^T) = \text{tr}((\rho_X W_X)(\rho_X W_X)^T) + \text{tr}((\rho W)(\rho W)^T) - 2\rho_X \rho \text{tr}(W_X W^T) = \|\rho_X W_X\|_F^2 + \|\rho W\|_F^2 - 2\rho_X \rho \text{tr}(W_X W^T)
\]

Standard matrix algebra permits the following identity to be verified for any quadratic form:

\[
x^T Ax = \text{tr}(Axx^T).
\]

Thus, the trace \(\text{tr}(W_X^T W)\) can be seen as the expected value of the quadratic form \(u^T W_X^T W u\), where \(u \in \mathbb{R}^n\) is a random vector such that \(\mathbb{E}(u) = 0 \in \mathbb{R}^n\) and \(\mathbb{E}(uu^T) = I_n\).
3 Proof of Equations (10) and (11)

Proof. In order to prove Equations (10) and (11), some preliminary results are first derived.

Since both $W_{(m)}$ and $W_{(\lfloor om \rfloor)}$ are binary matrices, we have that:

$$W_{(m)}\epsilon_n = m, \quad W_{(\lfloor om \rfloor)}\epsilon_n = \lfloor om \rfloor, \quad \text{(SM.4)}$$

where $\epsilon_n = [1, \ldots, 1]^T \in \mathbb{R}^n$.

The perturbation Equation (9) in matrix form becomes:

$$\tilde{W}_\alpha = A + [W_{(\lfloor om \rfloor)} - A] \odot B,$$

endequation where $\odot$ is the Hadamard matrix multiplication (that is, the elementwise multiplication). Just like in case of perturbation (9), $B_{ij} \sim B(1 - \gamma)$ for any $i \neq j$ and $B_{ii} = 0$ for any $i$, whereas the elements of $A$ are distributed as $A_{ij} \sim B(\lfloor am \rfloor, \lfloor (n-1) \rfloor)$ if $i \neq j$, whilst $A_{ii} = 0$ for any $i$. Off-diagonal elements of $A$ are statistically independent from off-diagonal elements of $B$.

Let define $Q_n \overset{\text{def}}{=} \eta_n - I_n \in \mathbb{R}^{n \times n}$. From the definition of $A$ and $B$, it follows that:

$$E(A) = (n - 1)^{-1} \lfloor am \rfloor \epsilon_n^\top \odot Q_n, \quad E(B) = (1 - \gamma)Q_n,$$

thus:

$$E(\tilde{W}_\alpha) = E(A) + W_{(\lfloor om \rfloor)} \odot E(B) - E(A) \odot E(B) =$$

$$= (n - 1)^{-1} \lfloor am \rfloor \epsilon_n^\top \odot Q_n + (1 - \gamma)W_{(\lfloor om \rfloor)} +$$

$$= (1 - \gamma)(n - 1)^{-1} \lfloor am \rfloor \epsilon_n^\top \odot Q_n + (1 - \gamma)W_{(\lfloor om \rfloor)}, \quad \text{(SM.5)}$$

since $Q_n \odot Q_n = Q_n$, and since for any spatial weight matrix $W$ of order $n$, $W \odot Q_n = W$.

Since both $W_{(m)}$ and $W_{(\lfloor om \rfloor)}$ are binary matrices based on the nearest-neighbour criterion, we have that:

$$(W_{(m)})_{ij} > 0 \quad \Rightarrow \quad (W_{(\lfloor om \rfloor)})_{ij} > 0 \quad \text{if } \alpha > 1$$

$$(W_{(\lfloor om \rfloor)})_{ij} > 0 \quad \Leftrightarrow \quad (W_{(m)})_{ij} > 0 \quad \text{if } \alpha = 1$$

$$(W_{(\lfloor om \rfloor)})_{ij} > 0 \quad \Rightarrow \quad (W_{(m)})_{ij} > 0 \quad \text{if } \alpha < 1$$

and thus:

$$W_{(m)} \odot W_{(\lfloor om \rfloor)} = W_{(\lfloor (1 \wedge \alpha) m \rfloor)}, \quad \text{(SM.6)}$$

Moreover, note that:

$$\epsilon^\top m = n\bar{m}, \quad m^\top m = n\bar{m}(1 + \kappa_m^2), \quad \text{(SM.7)}$$

by definition of $\bar{m} = \frac{1}{n}\sum_{i=1}^{n} m_i$ and $\kappa_m^2 = \frac{1}{nm^2} \sum_{i=1}^{n} (m_i - \bar{m})^2$.

Finally, define the following quantities:

$$d_\alpha \overset{\text{def}}{=} \alpha m - \lfloor am \rfloor, \quad \tilde{d}_\alpha = n^{-1}\epsilon^\top d_\alpha, \quad \text{(SM.8)}$$
and note that, since $0 \leq d_\alpha < \iota_n$, the following inequality holds:

$$0 \leq \tilde{d}_\alpha < 1 \quad \text{for any } \alpha \in \mathbb{R}^+$$  \hspace{1cm} (SM.9a)

whereas

$$\tilde{d}_\alpha = 0 \quad \text{for any } \alpha \in \mathbb{N}. \quad \text{(SM.9b)}$$

Correlation between the terms of $\tilde{W}_\alpha u$ and $W_{(m)} u$ requires covariance and variances to be computed. The rest of the proof is focused on this.

The expected values of the elements of $W_\alpha u$ and $W_{(m)} u$ is zero, thus the covariance between them can be computed as follows:

$$\mathbb{E} \left( \frac{1}{n} \sum_{i=1}^{n} (\tilde{W}_\alpha u)_i (W_{(m)} u)_i \right) = \mathbb{E} \left( \mathbb{E} \left( \frac{1}{n} \sum_{i=1}^{n} (\tilde{W}_\alpha u)_i (W_{(m)} u)_i \bigg| \tilde{W}_\alpha \right) \right) =$$

$$= n^{-1} \mathbb{E} \mathbb{E}(u^T \tilde{W}_\alpha^T W_{(m)} u | \tilde{W}_\alpha) =$$

$$= n^{-1} \mathbb{E} \mathbb{E}(\text{tr}(\tilde{W}_\alpha^T W_{(m)} \mathbb{E}(u u^T | \tilde{W}_\alpha))) =$$

$$= n^{-1} \mathbb{E} \mathbb{E}(\text{tr}(\tilde{W}_\alpha^T W_{(m)})) =$$

$$= n^{-1} \iota_n^T \mathbb{E}(\tilde{W}_\alpha \otimes W_{(m)}) \iota_n. \quad \text{(SM.10)}$$

From Equation (SM.6) it follows that:

$$\mathbb{E}(\tilde{W}_\alpha \otimes W_{(m)}) = \mathbb{E}(\tilde{W}_\alpha) \otimes W_{(m)} =$$

$$= \gamma (n-1)^{-1} [\alpha m] \iota_n^T \otimes W_{(m)} + (1-\gamma) W_{(1 \wedge \alpha) m} =$$

$$= \gamma (n-1)^{-1} \Lambda W_{(m)} + (1-\gamma) W_{(1 \wedge \alpha) m}, \quad \text{(SM.11)}$$

where $\Lambda \overset{\text{def}}{=} \text{diag}(\lfloor \alpha m \rfloor)$.

Hence, if (SM.11) is substituted in (SM.10), we have:

$$n^{-1} \iota_n^T \mathbb{E}(\tilde{W}_\alpha \otimes W_{(m)}) \iota_n =$$

$$= \gamma n^{-1} (n-1)^{-1} [\alpha m] \iota_n^T m + (1-\gamma) n^{-1} \iota_n^T \lfloor (1 \wedge \alpha) m \rfloor =$$

$$\approx \gamma (n-1)^{-1} (\alpha (1+\kappa_2^2) \bar{m}^2 - \tilde{m} d_\alpha) + (1-\gamma) \left( (1 \wedge \alpha) \bar{m} - \tilde{d}_\alpha \right),$$

because of (SM.4), and since

$$\iota_n^T \lfloor (1 \wedge \alpha) m \rfloor = (1 \wedge \alpha) \iota_n^T m - \iota_n^T \lfloor (1 \wedge \alpha) m - \lfloor (1 \wedge \alpha) m \rfloor \rfloor =$$

$$= (1 \wedge \alpha) \bar{m} \tilde{m} - n \tilde{d}_\alpha (1 \wedge \alpha), \quad \text{(SM.12a)}$$

$$[\alpha m]^T m = \alpha m^2 - (\alpha m - [\alpha m]^T m) =$$

$$= \alpha (\bar{m}^2 \kappa_2^2 + \bar{m}^2) - \tilde{d}_\alpha^2 \approx$$

$$\approx \alpha (1 + \kappa_2^2) \bar{m}^2 - \tilde{m} d_\alpha, \quad \text{(SM.12b)}$$

according to (SM.7).

Analogously, the variance of the elements of $\tilde{W}_\alpha u$ can be computed as follows:

$$\mathbb{E} \left( \frac{1}{n} \sum_{i=1}^{n} (\tilde{W}_\alpha u)_i^2 \right) = n^{-1} \iota_n^T \mathbb{E}(\tilde{W}_\alpha \otimes \tilde{W}_\alpha) \iota_n.$$

Note that $\tilde{W}_\alpha \otimes \tilde{W}_\alpha = W_\alpha$, since $\tilde{W}_\alpha$ is binary, hence:

$$\mathbb{E}(\tilde{W}_\alpha \otimes W_\alpha) = \mathbb{E}(W_\alpha)$$
and thus, according to Equation (SM.5) and (SM.12a), we have that:

\[ n^{-1}i_n E(W_\alpha \circ \tilde{W}_\alpha)_{tn} = n^{-1}i_n E(\tilde{W}_\alpha)_{tn} = \gamma n^{-1}i_n [\alpha m] + (1 - \gamma)n^{-1}i_n [\alpha m] = \alpha \bar{m} - \bar{d}_\alpha. \]

Finally, the variance of the elements of \( W_{(m)}u \) can be computed as it follows:

\[ E \left( \frac{1}{n} \sum_{i=1}^{n} (W_{(m)}u)^2_i \right) = n^{-1}i_n^2 (W_{(m)} \circ W_{(m)})_{tn} = \bar{m}. \]

It is now possible to determine the correlation coefficient between the elements of \( W_\alpha u \) and \( W_{(m)}u \) as it follows:

\[
\text{cor}(\tilde{W}_\alpha u, W_{(m)}u) = \frac{E \left( \frac{1}{n} \sum_{i=1}^{n} (\tilde{W}_\alpha u)_i (W_{(m)}u)_i \right)}{E \left( \frac{1}{n} \sum_{i=1}^{n} (\tilde{W}_\alpha u)_i^2 \right) \cdot E \left( \frac{1}{n} \sum_{i=1}^{n} (W_{(m)}u)_i^2 \right)} = \\
\gamma \frac{\bar{m}}{n-1} \alpha (1 + \kappa_m^2) - \frac{\bar{d}_\alpha}{m} \right] + (1 - \gamma) \left[ (1 \land \alpha) - \frac{\bar{d}_{1\land\alpha}}{m} \right] = \\
\sqrt{\alpha} \left( 1 - \frac{\bar{d}_\alpha}{am} \right)
\]

\[
= \gamma \frac{\bar{m}}{n-1} (1 + \kappa_m^2) \sqrt{\alpha} \xi_1 + (1 - \gamma) \frac{\sqrt{\alpha}}{\max \{1, \alpha\}} \xi_2, \quad \text{(SM.13)}
\]

where

\[
\xi_1 \equiv \frac{1 - \frac{\bar{d}_\alpha}{(1 + \kappa_m^2)am}}{\sqrt{1 - \frac{\bar{d}_\alpha}{am}}}, \quad \xi_2 \equiv \frac{1 - \frac{\bar{d}_{1\land\alpha}}{(1 \land \alpha)m}}{\sqrt{1 - \frac{\bar{d}_\alpha}{am}}}. \quad \text{(SM.14)}
\]

If \( \bar{d}_\alpha = 0 \) then \( \xi_1 = 1 \) and \( \xi_2 = 1 \), and Equation (SM.13) becomes Equation (11). Moreover, if \( \kappa_m^2 = 0 \), Equation (10) is obtained.

Properties (SM.9) permit one to verify that \( \bar{d}_{1\land\alpha} = 0 \) if \( \alpha \geq 1 \), hence \( \bar{d}_{1\land\alpha} = I_{\{\alpha < 1\}} \bar{d}_{1\land\alpha} \) (being \( I_{\{\cdot\}} \) the indicator function). Basic algebra manipulations permit the following inequalities to be derived:

\[
1 \leq \xi_1 \leq 1 + \frac{\kappa_m^2}{1 + \kappa_m^2} (\alpha \bar{m} - 1)^{-1},
\]

\[
1 - I_{\{\alpha < 1\}} (\alpha \bar{m})^{-1} \leq \xi_2 \leq 1 + I_{\{\alpha \geq 1\}} (\alpha \bar{m} - 1)^{-1},
\]

for \( \alpha \bar{m} > 1 \), whereas the condition \( \alpha \bar{m} \leq 1 \) is practically irrelevant. \( \square \)

## 4 Proof of Equation (13)

**Proof.** If the binary matrix \( H \in \{0, 1\}^{n \times n} \) is defined as:

\[ H_{ij} = I_{\{(W)_{ij} > 0\}}, \]

perturbation (12) can be restated in matrix form as it follows:

\[ \tilde{W} = V \odot (Q_n - B) \odot H + W \odot B, \quad \text{(SM.15)} \]
where $Q_n = \iota_n^T \iota_n - I_n$.

From (SM.15), it follows that:

$$
\bar{W} \odot \bar{W} = V \odot V \odot (Q_n - B) + W \odot W \odot B,
$$

$$
\tilde{W} \odot W = V \odot W \odot (Q_n - B) + W \odot W \odot B,
$$

since $(Q_n - B) \odot B = 0$. Thus:

$$
\mathbb{E}((\tilde{W} \odot \tilde{W})) = \gamma (1 + \kappa_V^2) \mu_V H + (1 - \gamma)(1 + \kappa_W^2) \mu_W H,
$$

$$
\mathbb{E}(\tilde{W} \odot W) = \gamma (1 + \kappa_W \kappa_V \rho_{WV}) \mu_W \mu_V H + (1 - \gamma)(1 + \kappa_W^2) \mu_W H,
$$

$$
\mathbb{E}(W \odot W) = (1 + \kappa_W^2) \mu_W^2 H.
$$

Finally we have that:

$$
\text{cor}(\tilde{W}_u, W_u) = \frac{1}{n} \frac{1}{\iota_n^T \iota_n} \mathbb{E}((\tilde{W} \odot W)_{\iota_n}) = \frac{\gamma (1 + \kappa_W \kappa_V \rho_{WV}) \mu_W \mu_V + (1 - \gamma)(1 + \kappa_W^2) \mu_W}{\sqrt{(1 + \kappa_W^2) \mu_W^2 \gamma (1 + \kappa_V^2) \eta^2 + (1 - \gamma)(1 + \kappa_W^2) \mu_W}}.
$$

This completes the proof. \hfill \Box

5 Proof of Equations (14) and (15)

5.1 Proof of Equation (15a)

Proof. For the sake of notational convenience, define the following quantities:

$$
a_W = 1 + \kappa_V^2, \quad a_V = 1 + \kappa_V^2, \quad a_{WV} = 1 + \kappa_W \kappa_V \rho_{WV}, \quad (SM.16)
$$

and the function $g$ as it follows:

$$
g(\gamma, \eta, a_W, a_V) = a_W \left[ \gamma a_V \eta^2 + (1 - \gamma) a_W \right];
$$

then note that:

$$
\frac{\partial g}{\partial \eta} = 2 \gamma a_W a_V \eta.
$$

It is now possible to compute the first derivative of (13) with respect to $\eta$ as it follows:

$$
\frac{\partial}{\partial \eta} \text{cor}(\tilde{W}_{\bar{u}}, W_{\bar{u}}) = \frac{\partial}{\partial \eta} \left( \frac{\gamma a_W a_V \eta + (1 - \gamma)a_W}{\sqrt{g(\gamma, \eta, a_W, a_V)}} \right) = \frac{\gamma a_W V g(\gamma, \eta, a_W, a_V) - [g(\gamma, \eta, a_W, a_V)]^{3/2} a_W a_V \eta}{[g(\gamma, \eta, a_W, a_V)]^{3/2}} = \frac{\gamma (1 - \gamma) a_W^2 (a_W V - a_V \eta)}{(a_W \left[ \gamma a_V \eta^2 + (1 - \gamma) a_W \right])^{3/2}}.
$$

If the previous derivative is set to zero and the equation is solved with respect to $\eta$, optimality condition (15a) is found. \hfill \Box
5.2 Proof of Equation (15b)

Proof. Using notation shortcuts defined in (SM.16), it is possible to compute the first derivative of (13) with respect to $\kappa_V$ as it follows:

$$\frac{d}{d\kappa_V} \text{cor}(\tilde{W}_m u, W_m u) = \frac{d}{d\kappa_V} \left( \frac{\gamma a_{WV} \eta + (1 - \gamma)a_W}{\sqrt{g(\gamma, \eta, a_W, a_V)}} \right) =$$

$$= \gamma \eta a_W \rho_{WV\kappa_W} \left[ \gamma a_V \eta^2 + (1 - \gamma)a_W \right] - \left[ \gamma a_{WV} \eta + (1 - \gamma)a_W \right] \kappa_V \eta =$$

$$= \gamma \eta (1 + \kappa_W^2) \frac{\gamma^2 (\rho_{WV\kappa_W} - \kappa_V) + (1 - \gamma)(1 + \kappa_W^2)(\rho_{WV\kappa_W} - \kappa_V \eta)}{[g(\gamma, \eta, \kappa_W, \kappa_W)]^{3/2}}.$$

If the previous derivative is set to zero and the equation is solved with respect to $\kappa_V$, optimality condition (15b) is found. \qed

5.3 Proof of Equation (14)

Proof. Correlation (13) is maximised with respect to $\eta$ and $\kappa_V$ if both conditions (15) are satisfied. It follows that if the system of two equations (15) is solved with respect to $\eta$ and $\kappa_V$, solution (14) is found.

This can be easily verified if (15a) is substituted in (13), and the result is maximised with respect to $\kappa_V$:

$$\frac{d}{d\kappa_V} \left( \frac{\gamma a_{WV} \eta + (1 - \gamma)a_W}{\sqrt{g(\gamma, \eta, a_W, a_V)}} \right) =$$

$$= \frac{d}{d\kappa_V} \left( \frac{\gamma a_{WV}^2 a_V^{-1} + (1 - \gamma)a_W}{\sqrt{a_W \left[ \gamma a_{WV}^2 a_V^{-1} + (1 - \gamma)a_W \right]}} \right)$$

$$= \frac{d}{d\kappa_V} \left( \sqrt{1 - \gamma + \frac{a_{WV}^2}{a_W a_V}} \right) =$$

$$= \left( 1 - \gamma + \frac{a_{WV}^2}{a_W a_V} \right)^{-1/2} \gamma \frac{a_{WV}}{a_W a_V} \left( \kappa_W \rho_{WV} - \kappa_V \right).$$

If the previous derivative is set to zero and the equation is solved with respect to $\kappa_V$, optimality condition (14) is found for $\kappa_V$. If $\kappa_V^* = \kappa_W \rho_{WV}$ is substituted in (15a), optimality condition $\eta^* = 1$ in (14) is found. \qed

References
