Online supplement for: Adding Spatially-Correlated Errors Can Mess Up The Fixed Effect You Love

James S. Hodges[1*], **Brian J. Reich**[2]

[1]Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota USA 55414
[2]Department of Statistics, North Carolina State University, Raleigh, North Carolina USA 27695
[*] *email:* hodge003@umn.edu

October 7, 2010

Part A: A small exploration of spatial confounding in some common geostatistical models

In the main paper's Section 2.5, each municipality i was assigned north-south and east-west coordinates. In the computations below, each coordinate was centered so it averaged zero across the municipalities, and both centered coordinates were divided by a common scaling constant. The symbols N_i and A_i refer to the centered and scaled north-south and east-west coordinates respectively, which had standard deviations 0.76 and 1.24 respectively.

We considered four forms of cov($\mathbf{S}$) determined by combinations of two distance measures and two specifications of correlation between municipalities as a function of distance between them. The measures describing distance between municipalities i and j, δ_{ij}, were Euclidean distance $[(N_i - N_j)^2 + (A_i - A_j)^2]^{0.5}$ and maximum distance $\max(|N_i - N_j|, |A_i - A_j|)$, for which the largest distances between two municipalities were 5.4 and 4.7 respectively. The forms specifying correlation as a function of distance were exponential, $\exp(-\delta_{ij}/\theta)$, and linear, $\max[(1 - \delta_{ij}/\theta), 0]$.

For each form of cov($\mathbf{S}$), Table 1 shows Pearson's correlation between the eigenvector corresponding to the j^{th} smallest eigenvalue of cov($\mathbf{S}$)$^{-1}$ and the eigenvector corresponding to the j^{th} smallest eigenvalue of the ICAR model's neighbor matrix $\mathbf{Q}$, for a range of values of cov($\mathbf{S}$)'s tuning constant θ. When this correlation is high for $j = 2$, the geostatistical specification for cov($\mathbf{S}$) will produce the same spatial confounding as the ICAR model.

Table 1 shows that generally the correlations between eigenvectors of cov($\mathbf{S}$)$^{-1}$ and $\mathbf{Q}$ are high for $j = 2$ and $j = 3$, but fall off substantially for $j = 4$ and for $j > 4$ (not shown). Results for the four cov($\mathbf{S}$) differ somewhat in details. The two distance measures behave similarly. However, while the exponential function of distance gives high correlations for all values of θ shown here for $j = 2, 3$, the linear function of distance shows smaller correlations for small values of θ, most likely because the spatial correlation dies off so quickly for small θ, and for $j = 3$ for large θ as well.

Table 1: Correlation between eigenvectors corresponding to j^{th} smallest eigenvalues of the ICAR's $\mathbf{Q}$ matrix and of $\text{cov}(\mathbf{S})^{-1}$, for four specifications of $\text{cov}(\mathbf{S})$ and various θ.

Distance	Correlation		θ									
measure	function	j	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
Euclidean	exponential	2	0.89	0.95	0.95	0.96	0.96	0.96	0.96	0.96	0.96	0.96
		3	0.79	0.87	0.88	0.88	0.88	0.87	0.87	0.86	0.86	0.85
		4	0.42	0.20	0.21	0.23	0.24	0.25	0.25	0.26	0.26	0.26
Euclidean	linear	2	0.34	0.61	0.88	0.92	0.94	0.95	0.96	0.96	0.96	0.96
		3	0.44	0.60	0.80	0.87	0.88	0.85	0.66	0.28	0.18	0.19
		4	0.25	0.33	0.10	0.15	0.21	0.27	0.28	0.20	0.15	0.13
maximum	exponential	2	0.90	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
		3	0.81	0.86	0.87	0.87	0.87	0.86	0.86	0.85	0.84	0.83
		4	0.17	0.11	0.10	0.10	0.09	0.09	0.09	0.08	0.08	0.08
maximum	linear	2	0.36	0.72	0.90	0.93	0.94	0.94	0.95	0.95	0.95	0.95
		3	0.15	0.66	0.83	0.87	0.87	0.87	0.26	0.25	0.26	0.29
		4	0.26	0.36	0.09	0.11	0.14	0.13	0.01	0.03	0.04	0.04

Part B: Spatial confounding with spectral methods

Section 2.2 of the main paper gives results for a discrete spatial domain; these can be extended to Gaussian process models defined on a continuous spatial domain. Let $y(\mathrm{s}_i) = \mathrm{x}_i'\boldsymbol{\beta} + S(\mathrm{s}_i) + \varepsilon_i$, where $\mathrm{s}_i \in \mathcal{R}^2$, $\varepsilon_i \overset{iid}{\sim} \mathrm{N}(0, \tau_e)$ is pure error, and S is a spatial process with mean zero, precision τ_s, and stationary spatial correlation function $\text{Cor}(\mathrm{s}_i, \mathrm{s}_j) = \rho(\mathrm{s}_i - \mathrm{s}_j)$.

The spectral representation theorem states that $S(\mathrm{s})$ can be written

$$S(\mathrm{s}) = \int_{\mathcal{R}^2} \cos(\boldsymbol{\omega}'\mathrm{s}) db_1(\boldsymbol{\omega}) + \int_{\mathcal{R}^2} \sin(\boldsymbol{\omega}'\mathrm{s}) db_2(\boldsymbol{\omega}), \tag{1}$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2)' \in \mathcal{R}^2$ is a frequency and b_1 and b_2 are independent Gaussian processes with mean zero, orthogonal increments, and $E(|db_j(\boldsymbol{\omega})|^2) = F(\boldsymbol{\omega})/\tau_s$. The spectral representation formulates the spatial process as a convolution of trigonometric basis functions and stochastic processes in the frequency domain with independent increments. The spatial correlation ρ is directly related to the spectral density F:

$$\rho(\mathrm{s}_i - \mathrm{s}_j) = \int_{\mathcal{R}^2} \cos[\boldsymbol{\omega}'(\mathrm{s}_i - \mathrm{s}_j)] dF(\boldsymbol{\omega}). \tag{2}$$

The spectral density is often decreasing in $||\boldsymbol{\omega}||$, for example, $F(\boldsymbol{\omega}) \propto \exp(-||\boldsymbol{\omega}||^2/(4\phi))$ corresponds to the squared-exponential covariance $\rho(\mathrm{s}_i - \mathrm{s}_j) = \exp(-\phi||\mathrm{s}_i - \mathrm{s}_j||^2)$.

Assuming the observations lie on a $m \times m$ square grid with distance one between neighboring sites, S's density can be approximated (Whittle 1954) by a sum over a finite grid of Fourier

frequencies $\boldsymbol{\omega}_l \in \{2\pi\lfloor -(m-1)/2 \rfloor/m, ..., 2\pi(m - \lfloor m/2 \rfloor)/m\}^2$,

$$\begin{aligned} S(\mathrm{s}) &= \sum_{j=1}^{2}\sum_{l=1}^{m^2} Z_j(\mathrm{s}, \boldsymbol{\omega}_l) b_{jl} \\ b_{jl} &\sim \mathrm{N}(0, \tau_s d_l), \end{aligned} \tag{3}$$

where $\lfloor x \rfloor$ is the smallest integer greater than or equal to x, $Z_1(\mathrm{s}, \boldsymbol{\omega}) = \cos(\boldsymbol{\omega}'\mathrm{s})$, $Z_2(\mathrm{s}, \boldsymbol{\omega}) = \sin(\boldsymbol{\omega}'\mathrm{s})$, and $\tau_s d_l$ is a precision with $d_l = 1/F(\boldsymbol{\omega}_l)$. In this special case, the Z_j are orthogonal, i.e., $\sum_{l=1}^{m^2} Z_j(\mathrm{s}, \boldsymbol{\omega}_l) Z_k(\mathrm{s}', \boldsymbol{\omega}_l) = I(k = j) I(\mathrm{s} = \mathrm{s}')$. This approximation may induce edge and aliasing effects in the spatial covariance, especially for small grids, but the approximation is useful for studying the fixed effects. This representation is directly analogous to the ICAR model as in equation (5) in the main paper, so the rest of the analysis applied to the ICAR model follows directly.

Comparing the spectral representation (3) with the ICAR model in equation (5) in the main paper, the trigonometric functions Z_1 and Z_2 are analogous to Section 2.2's eigenvectors Z_j, and the inverse of the spectral density $1/F(\omega)$ is analogous to the eigenvalues d_j in Section 2.2. Unlike the eigenvectors and eigenvalues of the ICAR model, Z_1, Z_2 and F have explicit forms, so the role of the covariate's spatial scale is more clear. High-frequency terms (large $||\boldsymbol{\omega}||$) have small prior variance (small $F(||\boldsymbol{\omega}||)$) and are shrunk a great deal, while low-frequency terms (small $||\boldsymbol{\omega}||$) have large prior variance (large $F(||\boldsymbol{\omega}||)$) and are shrunk relatively little. In this parameterization, the correlations between a covariate and the Z_j clearly describe the covariate's variation at different spatial scales.

References

Whittle P (1954). On stationary processes in the plane. *Biometrika*, **41**, 434-449.