THE TORUS STRUCTURE APPROXIMATION FOR THE SPATIAL AUTOCOVARIANCES

Winson Taam, Oakland University, Rochester, MI 48309
Brian S. Yandell, University of Wisconsin-Madison, Madison, WI 53706

Key Words and Phrases: Simultaneous autoregressive; conditional autoregressive; Kronecker product; circulant matrix; Fourier matrix.

1. INTRODUCTION

A torus neighborhood matrix W has been used extensively in spatial analysis (Besag and Moran, 1975, Mardia and Marshall, 1984, Besag, 1977, Moran, 1973a). Until Martin (1986) the justification seemed to be merely a mathematical convenience.

In time series analysis, circular assumption provides an approximate diagonalization of the covariance matrix, Fuller (1976). Extension of Fuller's (1976) result to higher dimension is given by Martin (1986). Martin demonstrated the approximate diagonalization of spatial autocovariance matrix and Taam and Yandell (1987) provided the same result with a slightly different formulation of the covariance matrix.

The objective of this note is to discuss this approximation and its application in several aspects of spatial analysis in 2-dimension. In the next section, we introduce some definitions and notations. In section 3, the approximation is examined. Section 4 discusses the applications of this approximation.

2. DIAGONALIZATION OF NEIGHBORHOOD MATRICES

We express the circulant and non-circulant neighborhood matrices for rectangular lattices, $\mathbf{W}$ and $\mathbf{M}$, in terms of building blocks $\mathbf{B}_n$ and Π_n. While most of this work generalizes to more complicated models, we restrict attention to the one step rook case.

Let us define the $n \times n$ primitives

$$\mathbf{B}_n = \begin{bmatrix} 0 & \mathbf{I}_{n-1} \\ 0 & 0 \end{bmatrix} \text{ and } \Pi_n = \begin{bmatrix} 0 & \mathbf{I}_{n-1} \\ 1 & 0 \end{bmatrix} = \mathbf{B}_n + \mathbf{B}_n^{(n-1)T}, \tag{2.1}$$

where $\mathbf{I}_{n-1}$ is an $(n-1) \times (n-1)$ identity matrix. Note that Π_n is a circulant matrix (Davis 1979), with the following eigenvalue-eigenvector decomposition:

$$\Pi_n = \mathbf{P}\Phi_n\mathbf{P}^* \text{ and } \Pi_n^T = \mathbf{P}\Phi_n^*\mathbf{P}^* \tag{2.2}$$

with $*$ denoting conjugate transpose,
$\{\mathbf{P}\}_{jk} = n^{-\frac{1}{2}}\exp(2\pi i\, jk/n)$ an orthogonal Fourier matrix, and Φ_n diagonal with $\{\Phi_n\}_{kk} = \exp(2\pi i k/n)$ for $j, k = 0, 1, 2, \cdots, n-1$. Let us introduce the matrix functionals

$$\mathbf{J}(k,n) = \begin{Bmatrix} \mathbf{I}_n & \text{if } k = 0 \\ \mathbf{B}_n^k & \text{if } k > 0 \\ \mathbf{B}_n^{kT} & \text{if } k < 0 \end{Bmatrix} \text{ and}$$
$$\mathbf{F}(k,n) = \begin{Bmatrix} \mathbf{I}_n & \text{if } k = 0 \\ \Pi_n^k & \text{if } k > 0 \\ \Pi_n^{-kT} & \text{if } k < 0 \end{Bmatrix} \tag{2.3}$$

Note that $\mathbf{B}_n^n = 0$, $\Pi_n^n = \mathbf{I}_n, \Pi_n^{n-1} = \Pi_n$ and $\Pi_n^j = \mathbf{B}_n^j + \mathbf{B}_n^{(n-j)T}$ for $j \leq n$. Therefore,

$$\mathbf{F}(x,n) = \begin{cases} \mathbf{J}(0,n) & \text{if } x = 0 \\ \mathbf{J}(x,n) + \mathbf{J}(x-n,n) & \text{if } x > 0 \\ \mathbf{J}(x,n) + \mathbf{J}(n-x,n) & \text{if } x < 0 \end{cases} \tag{2.4}$$

Let us define a lexicographic indexing system for an $r \times c$ two-dimensional lattice where we assign the numbers $j = 1, \cdots, N = rc$ to the elements (u,v) of the lattice, that is,

$$j(u,v) = 1 + cu + v, \quad u = 0, 1, \cdots, r-1,$$
$$v = 0, 1 \cdots, c-1, \tag{2.5}$$
$$\text{or } u_j = [(j-1)/c], \quad v_j = (j-1) \bmod c.$$

For simplicity, let us look at the "rook case" neighborhood matrix . We consider a rectangular lattice in which neighbors at the boundary are only those rook neighbors within the rectangle. This leads to unbalanced weights, which can be represented as the neighborhood matrix

$$\begin{aligned} \mathbf{M} &= \mathbf{I}_r \otimes (\mathbf{B}_c + \mathbf{B}_c^T) + (\mathbf{B}_r + \mathbf{B}_r^T) \otimes \mathbf{I}_c \\ &= \mathbf{J}(0,r) \otimes [\mathbf{J}(1,c) + \mathbf{J}(-1,c)] \\ &+ [\mathbf{J}(1,r) + \mathbf{J}(-1,r)] \otimes \mathbf{J}(0,c). \end{aligned} \tag{2.6}$$

This has an exact diagonalization (Conte and deBoor, 1980, p.206), based on the building blocks $\mathbf{B}_n + \mathbf{B}_n^T = \mathbf{Q}_n \Psi_n \mathbf{Q}_n^T$, where the orthogonal matrix $\mathbf{Q}_n$ and the diagonal matrix Ψ_n are defined as

$$\{\mathbf{Q}_n\}_{jk} = (2/(n+1))^{\frac{1}{2}}\sin(\pi jk/(n+1)) \text{ and}$$
$$\{\Psi_n\}_{kk} = 2\cos(\pi k/(n+1)), \; j, \; k = 1, \cdots, n.$$

The exact diagonalization $\mathbf{M}$ can be written as the Kronecker product $\mathbf{M} = (Q_r \otimes Q_C)[\mathbf{I}_r \otimes \Psi_C + \Psi_r \otimes \mathbf{I}_C](Q_r \otimes Q_c)^T = Q\Psi Q^T$ with Ψ diagonal, $\{\Psi\}_{kk} = 2[\cos(\pi u_k/(r+1)) + \cos(\pi v_k/(c+1))]$, $k = 1, \cdots, N$, and

$$\{\mathbf{Q}\}_{jk} = 2((r+1)(c+1))^{-\frac{1}{2}}\sin(\pi u_j u_k/(r+1))$$
$$\sin(\pi v_j v_k/(c+1)), \; j,k = 1, \cdots, N.$$

See also Besag's discussion of Bartlett (1978). Note, however, that $\mathbf{M}^k$ cannot be diagonalized by the same $\mathbf{Q}$ since there is no matrix that can diagonalize $\mathbf{B}_N^k + \mathbf{B}_N^{kT}$ for all $1 \leq k \leq N$. Such a diagonalization would prove useful for simplifying the computation of the spatial covariance. Therefore, it seems appropriate to consider approximate diagonalization of $\mathbf{M}$.

The circulant matrix which we use to approximate $\mathbf{M}$ is

$$\begin{aligned} \mathbf{W} &= \mathbf{I}_r \otimes (\Pi_c + \Pi_c^T) + (\Pi_r + \Pi_r^T) \otimes \mathbf{I}_c \\ &= \mathbf{F}(0,r) \otimes [\mathbf{F}(1,c) + \mathbf{F}(-1,c)] \\ &+ [\mathbf{F}(1,r) + \mathbf{F}(-1,r)] \otimes \mathbf{F}(0,c), \end{aligned} \tag{2.7}$$

which is diagonalized by $\mathbf{P}_N$,

$$\begin{aligned} \Phi_N &= \mathbf{P}_N \mathbf{W} \mathbf{P}_N^* \\ &= \mathbf{I}_r \otimes (\Phi_c + \Phi_c^*) + (\Phi_r + \Phi_r^*) \otimes \mathbf{I}_c \end{aligned} \tag{2.8}$$

with $\{\Phi_N\}_{kk} = 2[\cos(2\pi u_j/r) + \cos(2\pi v_k/c)]$. The orthogonal matrix $\mathbf{P}_N$ that diagonalizes $\mathbf{W}$ is the Kronecker product of the matrices that diagonalize the Π's. That is, $\mathbf{P}_N = \mathbf{P}_r \otimes \mathbf{P}_c$, with

$$\{\mathbf{P}_N\}_{jk} = (rc)^{-\frac{1}{2}}\exp(2\pi i((u_j u_k/r) + (v_j v_k/c))), \tag{2.9}$$

which is again a Fourier matrix. Note that

$$P_N(\Pi_r^j \otimes \Pi_c^k)P_N^* = (P_r \otimes P_c)(\Pi_r^j \otimes \Pi_c^k)(P_r^* \otimes P_c^*$$
$$= (P_r \Pi_r^j P_r^*) \otimes (P_c \Pi_c^k P_c^*)$$
$$= \Phi_r^j \otimes \Phi_c^k,$$

for all $j \leq m_r$ and $k \leq m_C$ with $m_r = [r/2]$ and $m_C = [c/2]$. Further, P_N diagonalizes Π_N^k for all k.

3. APPROXIMATE DIAGONALIZATION OF COVARIANCE IN 2 DIMENSION

We argue as in Fuller (1976) and Martin (1986) to show element-wise convergence of our approximate diagonalization of the covariance matrix for the 2-dimensional case. Since Whittle (1954) indicated that a one-dimensional bilateral scheme (two directional correlation) can be represented by a one-dimensional unilateral scheme, we concentrate on the two dimensional cases.

Suppose we have an $r \times c$ rectangular lattice with observations $\{Y_{t,s}\}$, $t = 0, 1, \cdots, r-1$ and $s = 0, 1, \cdots, c-1$ coming from some spatial process. We define the covariance matrix Γ as

$$\Gamma = \sum_{j=-(r-1)}^{r-1} \sum_{k=(c-1)}^{c-1} \gamma(j,k) J(j,r) \otimes J(k,c), \tag{3.1}$$

and its circular counterpart Γ_1 as

$$\Gamma_1 = \sum_{j=-m_r}^{m_r} \sum_{k=-m_c}^{m_c} \gamma(j,k) F(j,r) \otimes F(k,c) \tag{3.2}$$

where $\gamma(j,k) = \mathrm{Cov}(Y_{t,s}, Y_{t+j,s+k})$ and J and F are defined in (2.3). Note that the layout of Γ has not specified a spatial model. In addition, the $\gamma(j,k)$ here is not the same as γ_g^T of Martin (86) but Γ_1 is equivalent to Martin's $V_T = \sum_{g=-M}^{N} \gamma_g^T \left\{ \bigotimes_{u=1}^{2} W_{n_u}^{g_u} \right\}$ The 2-dimensional Fourier matrix P_N defined in (2.9) diagonalizes Γ_1 as

$$\Lambda = P_N^* \Gamma_1 P_N = \sum_{j=-m_r}^{m_r} \sum_{k=-m_c}^{m_c} \gamma(j,k) \Phi_r^j \otimes \Phi_c^k \tag{3.3}$$

where $\Lambda_{\ell\ell} = \sum_{j=-m_r}^{m_r} \sum_{k=-m_c}^{m_c} \gamma(j,k)$ $exp(2\pi i(ju_\ell/r + kV_\ell/c))$ and $\ell = 1, 2, \cdots N$. The approximation of Γ by Γ_1 is shown in the following theorem.

Theorem 1. Suppose that we have a rectangular $r \times c$ lattice from a stationary process and suppose that $\gamma(j,k) = \gamma(-j,-k)$ for $k, j = 0, \pm 1, \pm 2, \cdots$. Suppose that γ is absolutely summable with respect to both indices, $\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} |\gamma(k,j)| = g < \infty$. Let P_N, Γ, Γ_1 and Λ be defined as in (2.9), (3.1), (3.2), and (3.3) respectively. Then each element of $|P_N^*(\Gamma_1 - \Gamma)P_N| \to 0$ as $\min(r,c) \to \infty$, where $|A|$ is the matrix A with entries equal to their absolute values.

The proof of this result which is given in the Appendix is similar to the proof of theorem 4.2.1 in Fuller (1976) and the one in Martin (86). The approximation rate depends on the grid size as seen in the following:

Corollary 2. Under the same assumptions as in Theorem 1, the convergent rate of the elements of the absolute difference between Γ and Γ_1 is of $O(1/q)$ with $q = \min(r,c)$.

The proof of this corollary is given in the Appendix as well.

Remark There is no restriction that the spatial models be autoregressive or moving average, conditional or simultaneous. The general assumption needed is the summability of the covariance function.

4. Applications

4.1 Boundary Problem

In time series, there are three ways to handle the boundary problem. One approach is to assume a circular structure on the observed series. The other two approaches are the conditional likelihood and exact likelihood. In the current note, we discuss only the circular (torus) approximation for lattice data. For further discussion on boundary problems in spatial data, see Griffith (1983, 1985) and Martin (1987).

Assumption on a torus structure not only provides a good approximation of the covariance but also avoids the boundary problem in modeling autoregressive moving average spatial process.

When one considers modelling spatial processes on an infinite size grid, there are two groups of autoregressive moving average models. One uses the conditional argument, Besag (1974,1977), Besag and Moran (1975), and Moran (1973a,1973b). Its basic first order autoregressive model is defined as

$$E(Y_i | \text{all } Y\text{'s}) = \rho \sum_{j \in N(i)} Y_j \tag{4.1}$$

where Y_j represents the response at the j^{th} location and ρ is the first order autoregressive parameter. $N(i)$ is the index set associated with the immediate neighbors of i.

Another group uses the simultaneous argument, Whittle (1954) and Tjostheim (1978, 1983). Its basic first order autoregressive model is

$$Y_i = \rho \sum_{j \in N(i)} Y_j + \epsilon_i. \tag{4.2}$$

Again, $N(i)$ is the index set associated with the immediate neighbors of i and ϵ_i is a white noise process. These two groups of models rely on the fact that these processes come from an infinite lattice. When one is given a finite rectangular lattice, the parameter estimation requires the exact form of the covariance. Often times there is no easy closed form of the covariance function for these spatial models. Hence, the idea of using a circular structure (torus) and letting the size of a torus go to infinity has been discussed in Besag and Moran (1975) and Moran (1973a). We have used that idea and derived the exact expression for the covariance function $\gamma(j,k)$ under the first order autoregressive processes. Besag (81) obtained equivalent expression through autocovariance generating function. This fits in with our discussion of approximate diagonalization of spatial covariance matrix. Moran, Besag and others have also pointed out that when the size of the torus goes to infinity, the spectral representation of this process gives the same representation of the infinite lattice process. Therefore, we use this concept to show that a simple covariance matrix denoted by Γ_2 soon to be defined can approximate Γ_1, which in turn approximates Γ because of Theorem 1.

For simplicity, let us assume simple first order models and rook case immediate neighbor structure. Let us define Γ_2 as

$$\begin{aligned} \Gamma_2 &= \sigma^2(I - \rho W)^{-1} \text{ for a conditional } AR(1) \\ &= \sigma^2(I - \rho W)^{-2} \text{ for a simultaneous } AR(1) \end{aligned} \tag{4.3}$$

where W is defined in equation (2.7). Note that the entries in Γ_2 are not the same as the entries in Γ_1 because Γ_1 has a circular structure but its entries are the theoretical covariance, $\gamma(j,k)$, which can be expressed in spectral representation. The entries of Γ_2 however are the true $\gamma(j,k)$ plus some extra terms. Again the entries differ from γ_g^T of Martin (86). These extra terms are later showed to be negligible. We show that the covariance matrix Γ_2 for the first order process can be used to approximate Γ_1 in the following theorem.

Theorem 3. Assume that a process is conditional $AR(1)$. Suppose Γ_1 is defined in (3.2) and Γ_2 is defined in (4.3). Let $0 < \rho < 1/4$, then each element of $|\Gamma_1 - \Gamma_2| \to 0$ as $\min(r,c) \to \infty$.

The proof is given in the Appendix.

Remark: Using triangular inequality, this theorem implies that we can use Γ_2 to approximate Γ when the lattice is large. Note that the form of Γ_2 is very simple and its eigenvalues and eigenvectors are readily available.

Recall the approximate covariance matrix of the first order simultaneous autoregressive process $SAR(1)$ given in (4.3) is $\Gamma_2 = \sigma^2(\mathbf{I}-\rho\mathbf{W})^{-2}$.

Corollary 4. If the process is a first order simultaneous autoregressive, $SAR(1)$, Theorem 3 also holds.

The proof is essentially the same that of Theorem 3.

Caution:

Finally, let us examine one more definition of a covariance matrix. Cliff and Ord (1981, p.148) and Ripley (1981, p.89) define the covariance of the spatial process on a lattice without an explicit treatment on the boundary and use a neighborhood matrix without specifying the weights for the boundary plots.

Let us denote the covariance of a rectangular lattice with truncated boundary as Γ_3. That is

$$\begin{aligned}\Gamma_3 &= \sigma^2(\mathbf{I}-\rho\mathbf{M})^{-1} \text{ for } CAR(1)\\ &= \sigma^2(\mathbf{I}-\rho\mathbf{M}^{-T}(\mathbf{I}-\rho\mathbf{M})^{-1} \text{ for } SAR(1)\end{aligned}$$

where $\mathbf{M}$ is defined in (2.6). This matrix Γ_3 is used in likelihood estimation in Ord (1975) and in Besag's and Ripley's discussions of Bartlett (1978). However, one should be careful with the use of these two covariance matrices. In the $CAR(1)$ model, the difference

$$\begin{aligned}&\sigma^{-2}(\Gamma_2-\Gamma_3)\\ &= (\mathbf{I}-\rho\mathbf{W})^{-1}-(\mathbf{I}-\rho\mathbf{M})^{-1}\\ &= (\mathbf{I}-\rho\mathbf{W})^{-1}[\rho(\mathbf{W}-\mathbf{M})+\rho^2(\mathbf{WM}-\mathbf{M}^2)\\ &\quad+\rho^3(\mathbf{WM}^2-\mathbf{M}^3)+\cdots]\\ &= (\mathbf{I}-\rho\mathbf{W})^{-1}[\rho(\mathbf{W}-\mathbf{M})](\mathbf{I}-\rho\mathbf{M})^{-1}.\end{aligned}$$

The term in the square brackets does not vanish no matter how large r, c are. In fact, each element of $\sigma^{-2}|\Gamma_2-\Gamma_3|-|\rho||\mathbf{W}-\mathbf{M}|$ is positive. That is each element of $\sigma^{-2}|\Gamma_2-\Gamma_3| \geq \mathbf{G}(r,c,\rho) > 0$ for any (r,c) and $\rho > 0$. By the triangular inequality, one can see that each element of the absolute difference $|\Gamma_1-\Gamma_2|-|\Gamma_2-\Gamma_3|$ and $|\Gamma-\Gamma_3|-|\Gamma_2-\Gamma_3|$ are bounded below by some positive constant for whatever size lattice. Hence, Γ_3 is not a good approximation to the theoretical covariance. Furthermore, the process determined by this covariance (Γ_3) has not been identified in the literature. One possible interpretation is to consider it as a process that comes from a $r \times c$ lattice bounded with zero on all sides. In example 4 of Guyon (1982), Guyon has pointed out the inconsistency of estimating autocorrelation under this covariance structure (Γ_3).

Remark: Numerical comparison of Γ and Γ_2 with $SAR(1)$ and $CAR(1)$ models show that for moderate sized lattices (11×11) and moderate correlation ($\rho = 0.1$), the approximate $\gamma(j,k)$ values are quite close (accurate to 5 decimal places). For somewhat smaller grids (8×8), the approximation is good (3-5 decimal places) for $j,k = 0,1,2,3$. The approximation may even be useful for grids as small as (6×6) if one is only interested in a few decimal places for $j,k = 0,1,2$. When $\rho = 0.24$ which is near its upper limit for stationarity, the approximation is not very good even for large grids. Torus structure is not meant for small lattices.

4.2 Expression for $\gamma(j,k)$

Expressions of $\gamma(j,k)$ for conditional and simultaneous $AR(1)$ given by Besag (81) can be obtained using a torus structure.

From Moran (1973a, 1973b), the covariance of conditional $AR(1)$ on an $r\times c$ grid with different parameters for the vertical and horizontal correlations is expressed as

$$\gamma(k,j) = \frac{\sigma^2}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi}\frac{cos(k\theta_1)cos(j\theta_2)d\theta_1 d\theta_2}{1-2\rho_1 cos(\theta_1) - 2\rho_2 cos(\theta_2)}$$

where ρ_1 measures the row (east-west) autocorrelation and ρ_2 measures the column (north-south) autocorrelation.

Note that $\gamma(j,0) \neq \gamma(0,j)$ if $\rho_1 \neq \rho_2$.

Using the spectral representation given in Whittle (1954), the covariance of simultaneous $AR(1)$ with one parameter is defined as

$$\gamma(k,j) = \frac{\sigma^2}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi}\frac{cos(k\theta_1)cos(j\theta_2)\,d\theta_1 d\theta_2}{[1-2\rho(cos(\theta_1) + cos(\theta_2))]^2}$$

where the variance of the white noise is σ^2 and the autoregressive parameter is ρ.

Suppose that $\gamma(j,k)$ is the covariance function of a $CAR(1)$ process with one parameter ρ. It can be written in the following expression.

$$\begin{aligned}\gamma(j,k) &= \sigma^2\sum_{s=0}^{\infty}\rho^{2s+j+k}\sum_{m=0}^{s}\binom{2s+j+k}{2m+j}\binom{2m+j}{m}\binom{2s-2m+k}{s-m}\\ &= \sigma^2\sum_{s=0}^{\infty}\rho^{2s+j+k}\binom{2s+j+k}{s+j}\binom{2s+j+k}{s}.\end{aligned} \tag{4.4}$$

This is obtained by expanding $(\mathbf{I}-\rho\mathbf{W}_\infty)^{-1}$ and collecting the terms of $\Pi_\infty^j \otimes \Pi_\infty^k$, where the subscript "$\infty$" means that the matrix is of infinite size. One can consider a $r\times c$ torus and let r and c go to infinity.

$$\begin{aligned}(\mathbf{I}-\rho\mathbf{W}_\infty)^{-1} &= \lim_{r,c\to\infty}(\mathbf{I}-\rho\mathbf{W}_N)^{-1}\\ &= \lim_{r,c\to\infty}\sum_{k=1}^{\infty}\rho^k(\mathbf{W}_N)^k\end{aligned} \tag{4.5}$$

where $N = rc$. Recall that $\mathbf{W}_N = \mathbf{I}_r\otimes(\Pi_c+\Pi_c^T)+(\Pi_r+\Pi_r^T)\otimes\mathbf{I}_c$. Hence,

$$\begin{aligned}(\mathbf{W}_N)^2 &= \mathbf{I}_r\otimes(\Pi_c^2+\Pi_c^{2T}+2\mathbf{I}_c) + 2(\Pi_r+\Pi_r^T)\\ &\quad\otimes(\Pi_c+\Pi_c^T) + (\Pi_r^2+\Pi_r^{2T}+2\mathbf{I}_r)\otimes\mathbf{I}_c,\\ (\mathbf{W}_N)^3 &= \mathbf{I}_r\otimes[\Pi_c^3+\Pi_c^{3T}+3(\Pi_c+\Pi_c^T)]\\ &\quad+3(\Pi_r+\Pi_r^T)\otimes(\Pi_c+\Pi_c^T)^2\\ &\quad+3(\Pi_r+\Pi_r^T)^2\otimes(\Pi_c+\Pi_c^T)\\ &\quad+[\Pi_r^3+\Pi_r^{3T}+3(\Pi_r+\Pi_r^T)]\otimes\mathbf{I}_c\end{aligned}$$

and so on. Then one collects the terms in the infinite sum of (4.5) which leads to

$$\begin{aligned}&\mathbf{I}_N(1+4\rho^2 + 36\rho^4 + \cdots)\\ &\mathbf{I}_r\otimes(\Pi_c+\Pi_c^T)(\rho+9\rho^3 + \cdots)\\ &(\Pi_r+\Pi_r^T)\otimes\mathbf{I}_c(\rho+9\rho^3+\cdots) \text{ etc.}\end{aligned}$$

Since we are dealing with an infinite torus, the γ function is given above in(4.4). Note that

$$\sigma^2\left[(\mathbf{I}-\rho\mathbf{W}_\infty)^{-1}\right]_{rc\times rc} = \Gamma$$

i.e., the $rc\times rc$ submatrix is the theoretical covariance matrix of a $CAR(1)$ process on a $r\times c$ rectangular lattice.

If we follow the same argument, we can obtain the covariance function for the first order simultaneous autoregressive process.

$$
\begin{aligned}
\gamma(j,k) &= \sigma^2 \sum_{s=0}^{\infty} (2s+j+k+1)\rho^{2s+j+k} \sum_{m=0}^{s} \binom{2s+j+k}{2m+j} \binom{2m+j}{m} \binom{2s-2m+k}{s-m} \\
&= \sigma^2 \sum_{s=0}^{\infty} (2s+j+k+1)\rho^{2s+j+k} \binom{2s+j+k}{s+j} \binom{2s+j+k}{s}.
\end{aligned}
$$

This is obtained by expanding $(\mathbf{I} - \rho \mathbf{W}_\infty)^{-2}$ and collecting the terms. Similarly, one can find an expression for $CAR(1)$ with two parameters as in the first example.

Remark: One advantage of these expressions is the saving in numerical evaluation of $\gamma(j,k)$. They require no numerical integration.

4.3 Applications in Field Trial Experiment

In field trial experiments, the use of this torus approximation in maximum likelihood estimation is common. For instance, Mardia and Marshall (1984), Besag (1974), Besag and Moran (1975) and Besag and Kempton (1986) used Γ_2 in one model or another.

One model is given by $Y = D\tau + \xi + \epsilon$ where Y is the response, D represents the design matrix and τ the corresponding effect, ϵ is the mean zero white noise process with variance σ_ϵ^2, and ξ is a random variable representing the spatial component with autocovariance $\gamma(j,k)$.

$$
\begin{aligned}
E(Y) &= D\tau \\
V(Y) &= V(\xi) + \sigma_\epsilon^2 I = \sigma_\epsilon^2 (I + \nu\Sigma)
\end{aligned}
$$

One can use Γ_2 to approximate $\sigma_\epsilon^2 \nu \Sigma$ and use maximum likelihood approach to estimate all the parameters.

4.4 Characterization of Spatial Process

In practice, the torus assumption seems unrealistic and uncanny. Nevertheless, the torus assumption provides a basis of the development of Markovian processes and the spectral representation of nearest neighbor processes discussed in Moran (1973a, 1973b), Bartlett (1971), Whittle (1954) and Besag (1974). Detail discussions on non-toroidal spatial processes are given by Tjostheim (1978, 1983, 1985).

APPENDIX

Prior to the proof of Theorem 1, let us break down the **F** and **J** notation in terms of the Π's and **B**'s.

$$
\begin{aligned}
\mathbf{W}_1^{(k)} &= \mathbf{I}_r \otimes (\Pi_c^k + \Pi_c^{kT}) \\
&= \mathbf{F}(0,r) \otimes (\mathbf{F}(k,c) + \mathbf{F}(-k,c)), \\
\mathbf{W}_2^{(j)} &= (\Pi_R^j + \Pi_r^{jT}) \otimes \mathbf{I}_c \\
&= (\mathbf{F}(j,r) + \mathbf{F}(-j,r)) \otimes \mathbf{F}(0,c) \\
\mathbf{W}_3^{(j,k)} &= (\Pi_r^j \otimes \Pi_c^k) + (\Pi_r^{jT} \otimes \Pi_c^{kT}) \\
&= (\mathbf{F}(j,r) \otimes \mathbf{F}(k,c)) + (\mathbf{F}(-j,r) \otimes \mathbf{F}(-k,c)), \\
\mathbf{W}_4^{(j,k)} &= (\Pi_r^j \otimes \Pi_c^{kT}) + (\Pi_r^{jT} \otimes \Pi_c^k) \\
&= (\mathbf{F}(j,r) \otimes \mathbf{F}(-k,c)) + (\mathbf{F}(-j,r) \otimes \mathbf{F}(k,c)),
\end{aligned}
$$

$\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ and $\mathbf{W}_4$ represent the horizontal, vertical, and the two diagonal neighborhoods, respectively. Let us similarly define the non-circular neighborhood matrices as **M**'s.

$$
\begin{aligned}
\mathbf{M}_1^{(k)} &= \mathbf{I}_r \otimes (\mathbf{B}_c^k + \mathbf{B}_c^{kT}), \\
\mathbf{M}_2^{(j)} &= (\mathbf{B}_r^j + \mathbf{B}_r^{jT}) \otimes \mathbf{I}_c, \\
\mathbf{M}_3^{(j,k)} &= (\mathbf{B}_r^j \otimes \mathbf{B}_c^k) + (\mathbf{B}_r^{jT} \otimes \mathbf{B}_c^{kT}) \text{ and} \\
\mathbf{M}_4^{(j,k)} &= (\mathbf{B}_r^j \otimes \mathbf{B}_c^{kT}) + (\mathbf{B}_r^{jT} \otimes \mathbf{B}_c^k).
\end{aligned}
$$

With the covariance being radially symmetric, $\gamma(j,k) = \gamma(-j,k)$ and $\gamma(-j,k) = \gamma(j,-k)$ for $j = 0, 1, \cdots, r-1$ and $k = 0, 1, \cdots, c-1$. Let us define

$$
\begin{aligned}
\gamma_1(k) &= \gamma(0,k), \\
\gamma_2(j) &= \gamma(j,0), \\
\gamma_3(j,k) &= \gamma(j,k) \text{ and} \\
\gamma_4(j,k) &= \gamma(j,-k).
\end{aligned} \qquad (A.1)
$$

Then (3.1) can be written as

$$
\begin{aligned}
\Gamma &= \gamma(0,0)\mathbf{I}_N \\
&+ \sum_{k=1}^{c-1} \gamma_1(k)\mathbf{M}_1^{(k)} \\
&+ \sum_{j=1}^{r-1} \gamma_2(j)\mathbf{M}_2^{(j)} \qquad (A.2) \\
&+ \sum_{j=1}^{r-1}\sum_{k=1}^{c-1} [\gamma_3(j,k)\mathbf{M}_3^{(j,k)} \\
&+ \gamma_4(j,k)\mathbf{M}_4^{(j,k)}],
\end{aligned}
$$

and similarly for Γ_1 of (3.2). With the assumption of radial symmetry, the diagonal matrix Λ in (3.3) becomes:

$$
\begin{aligned}
\Lambda &= \gamma(0,0)\mathbf{I}_N \\
&+ \sum_{k=1}^{m_c} \gamma_1(k)(\mathbf{I}_r \otimes (\Phi_c^k + \Phi_c^{K^*})) \\
&+ \sum_{j=1}^{m_r} \gamma_2(j)((\Phi_r^j + \Phi_r^{j^*}) \otimes \mathbf{I}_c) \\
&+ \sum_{k=1}^{c-1}\sum_{j=1}^{r-1} \{\gamma_3(j,k)[(\Phi_r^j \otimes \Phi_c^k) + (\Phi_r^{j^*} \otimes \Phi_c^{k^*})] \\
&+ \gamma_4(j,k)[(\Phi_r^{j^*} \otimes \Phi_c^k) + (\Phi_r^j \otimes \Phi_c^{k^*})]\},
\end{aligned}
$$

Let us now show the asymptotic diagonalization of Γ in (A.2) by $\mathbf{P}_N$ in (2.9).

Proof of Theorem 1. Without loss of generality, let us assume both r and c are odd. The proof is similar for either one or both of r and c even. Recall that $m_r = [r/2]$ and $m_c = [c/2]$.

$$
\begin{aligned}
\Gamma_1 - \Gamma = & \sum_{k=1}^{m_c} \gamma_1(k)\mathbf{W}_1^{(k)} - \sum_{k=1}^{c-1} \gamma_1(k)\mathbf{M}_1^{(k)} \\
& + \sum_{j=1}^{m_r} \gamma_2(j)\mathbf{W}_2^{(j)} - \sum_{j=1}^{r-1} \gamma_2(j)\mathbf{M}_2^{(j)} \\
& + \sum_{j=1}^{m_r}\sum_{k=1}^{m_c} [\gamma_3(j,k)\mathbf{W}_3^{(j,k)} + \gamma_4(j,k)\mathbf{W}_4^{(j,k)}] \\
& - \sum_{j=1}^{r-1}\sum_{k=1}^{c-1} [\gamma_3(j,k)\mathbf{M}_3^{(j,k)} + \gamma_4(j,k)\mathbf{M}_4^{(j,k)}] \\
= & \sum_{k=1}^{m_c} [\gamma_1(k) - \gamma_1(c-k)]\mathbf{M}_1^{(c-k)} \\
& + \sum_{j=1}^{m_r} [\gamma_2(j) - \gamma_2(r-j)]\mathbf{M}_2^{(r-j)} \\
& + \sum_{j=1}^{m_r}\sum_{k=1}^{m_c}\sum_{\ell=3}^{4} \{[\gamma_\ell(j,k) - \gamma_\ell(r-j,c-k)] \\
& \quad \mathbf{M}_\ell^{(r-j,c-k)} + [\gamma_\ell(j,k) - \gamma_\ell(r-j,k)]\mathbf{M}_\ell^{(r-j,k)} \\
& + [\gamma_\ell(j,k) - \gamma_\ell(j,c-k)]\mathbf{M}_\ell^{(j,c-k)}\}.
\end{aligned}
$$

Note that the diagonal elements of $\mathbf{P}_N^*(\Gamma_1 - \Gamma)\mathbf{P}_N$ are larger than the off-diagonal elements on the same row. That is

$$|\mathbf{P}_{\mu\cdot}^*(\Gamma_1 - \Gamma)\mathbf{P}_{\cdot\xi}| \le |\mathbf{P}_{\mu\cdot}^*(\Gamma_1 - \Gamma)\mathbf{P}_{\cdot\mu}| \text{ for } \mu \ne \xi.$$

Thus it suffices to show that $|\mathbf{P}_{\mu\cdot}^*(\Gamma_1 - \Gamma)\mathbf{P}_{\cdot\mu}| \to 0$ as r and $c \to \infty$ for each μ.

$$
\begin{aligned}
& |\mathbf{P}_{\mu\cdot}^*(\Gamma_1 - \Gamma)\mathbf{P}_{\cdot\mu}| \\
& = \Big| \sum_{k=1}^{m_c} [\gamma_1(k) - \gamma_1(c-k)] \frac{2k}{c} cos(\frac{2\pi}{c} k v_\mu) \\
& + \sum_{j=1}^{m_r} [\gamma_2(j) - \gamma_2(c-j)] \frac{2j}{r} cos(\frac{2\pi}{c} j u_\mu) \\
& + \sum_{j=1}^{m_r}\sum_{k=1}^{m_c} 2cos(2\pi((ju_\mu/r) + (kv_\mu/c)))/rc \\
& \times [(r-j)k(\gamma_4(j,k) - \gamma_4(j,c-k)) \\
& + j(c-k)(\gamma_4(j,k) - \gamma_4(r-j,k)) \\
& + jk(\gamma_3(j,k) - \gamma_3(r-j,c-k))] \\
& + 2cos(2\pi((ju_\mu/r) - (kv_\mu/c)))/rc \\
& \times [(r-j)k(\gamma_3(j,k) - \gamma_3(j,c-k)) \\
& + j(c-k)(\gamma_3(j,k) - \gamma_3(r-j,k)) \\
& + jk(\gamma_4(j,k) - \gamma_4(r-j,c-k))] \Big|
\end{aligned} \tag{A.4}
$$

$$
\begin{aligned}
\le (2/rc) \Big\{ & \sum_{k=1}^{m_c} rk|\gamma_1(k) - \gamma_1(c-k)| \\
& + \sum_{j=1}^{m_r} cj|\gamma_2(j) - \gamma_2(r-j)| \\
& + \sum_{j=1}^{m_r}\sum_{k=1}^{m_c}\sum_{\ell=3}^{4} [(r-j)k|\gamma_\ell(j,k) - \gamma_\ell(j,c-k)| \\
& + j(c-k)|\gamma_\ell(j,k) - \gamma_\ell(r-j,k)| \\
& + jk|\gamma_\ell(j,k) - \gamma_\ell(r-j,c-k)|] \Big\}
\end{aligned} \tag{A.5}
$$

To show the bound in (A.5) goes to zero, we need the following lemma.

LEMMA 5 (Two Dimensional Kronecker Lemma)

Let $\{a_i\}$ and $\{b_i\}$ be two increasing sequences of positive numbers. Let $\{X_{ij}\}$ be a lattice sequence of non-negative real numbers, which satisfy $\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} X_{ij} = X < \infty$, then $\sum_{i=1}^{r}\sum_{j=1}^{c} \frac{a_i}{a_r}\frac{b_j}{b_c} X_{ij} \to 0$, $\sum_{i=1}^{r} \frac{a_i}{a_r} X_{ij} \to 0$ for all j and $\sum_{j=1}^{c} \frac{b_j}{b_c} X_{ij} \to 0$ for all i as $\min(r,c) \to \infty$.

Proof of Lemma 5. Let $S_{r,c} = \sum_{i=1}^{r}\sum_{j=1}^{c} X_{ij}$ be the finite sums. Then

$$
\begin{aligned}
& \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{a_i}{a_r}\frac{b_j}{b_c} X_{ij} \\
& = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{a_i}{a_r}(S_{i,j} - S_{i,j-1} - S_{i-1,j} + S_{i-1,j-1}) \\
& = S_{r,c} - S_{r,c-1} - S_{r-1,c} + S_{r-1,c-1} \\
& + \sum_{j=1}^{c-1} \frac{b_j}{b_c}[S_{r,j} - S_{r-1,j} - S_{r,j-1} + S_{r-1,j-1}] \\
& + \sum_{i=1}^{r-1} \frac{a_i}{a_r}[S_{i,c} - S_{i,c-1} - S_{i-1,c} + S_{i-1,c-1}] \\
& + \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{a_i}{a_r}\frac{b_j}{b_c}(S_{i,j} - S_{i,j-1} - S_{i-1,j} + S_{i-1,j-1}) \\
& = S_{r,c} - S_{r,c-1} - S_{r-1,c} \\
& + \sum_{i=0}^{r-1}\sum_{j=0}^{c-1} \frac{(a_{i+1} - a_i)(b_{j+1} - b_j)}{a_r b_c} S_{i,j}.
\end{aligned} \tag{A.6}
$$

Note that $\{a_i\}$ and $\{b_i\}$ are increasing, so both $(a_{i+1} - a_i)$ and $(b_{j+1} - b_j)$ are positive. Without loss of generality, let $S_{0,j} = S_{i,0} = S_{0,0} = 0$ and $a_0 = b_0 = 0$. Since the $S_{i,j} \to X$ as $i, j \to \infty$, the expression in (A.6) goes to zero as $r, c \to \infty$. Because the X_{ij} are non-negative, the convergence of the marginal sums to zero follows.

By applying the above lemma with $X_{jk} = |\gamma(j,k)|$, it is obvious that $\sum_{k=1}^{m_c} 2k|\gamma_1(k) - \gamma_1(c-k)|/c$, $\sum_{j=1}^{m_r} 2j|\gamma_2(j) - \gamma_2(r-j)|/r$ and $\sum_{j=1}^{m_r}\sum_{k=1}^{m_c}\sum_{\ell=3}^{4} 2j(k)|\gamma_\ell(r-j,c-k)|/rc$ go to zero as $\min(r,c) \to \infty$ because of the absolute summability of γ. To show that

$$
\begin{aligned}
(2/rc) & \sum_{j=1}^{m_r}\sum_{k=1}^{m_c}\sum_{\ell=3}^{4} [k(r-j)|\gamma_\ell(j,k) - \gamma_\ell(j,c-k)| \\
& + j(c-k)|\gamma_\ell(j,k) - \gamma_\ell(r-j,k)|] \to 0,
\end{aligned}
$$

we examine one of these terms:

$$
\begin{aligned}
(2/rc) & \sum_{j=1}^{m_r}\sum_{k=1}^{m_c} [k(r-j)|\gamma_3(j,k) - \gamma_3(j,c-k)| \\
& \le (2/rc) \sum_{j=1}^{m_r}\sum_{k=1}^{m_c} (kr - jk)[|\gamma_3(j,k)| \\
& + |\gamma_3(j,c-k)|] \\
& \le (2k/c) \sum_{k=1}^{m_c} [g_{k\cdot} + g_(c-k)\cdot] - (e_1 + e_2)
\end{aligned}
$$

where $g_{k\cdot} = \sum_{j=1}^{\infty} |\gamma(j,k)| < \infty$, and e_1 and e_2 are the double summations with weights jk, which go to zero by Lemma 5. Applying the one-dimensional Kronecker Lemma to terms involving $g_{k\cdot}$ gives convergence to zero. Hence, the proof is completed for r and c odd.

Note that if either one or both of r and c is even, replace m_r or m_c in (A.4) and (A.5) by $m_r - 1$ or $m_c - 1$. Then the proof can still be carried through.

Proof of Corollary 2. From the absolute summability of γ, we have $|\gamma(k,j)| = O(1/rc)$. Let us examine the upper bound in (A.5)

$$\sum_{k=1}^{m_c} \frac{2k}{c} |\gamma_1(k) - \gamma_1(c-k)| = O(1/rc)$$

and

$$\sum_{j=1}^{m_r} \frac{2j}{r} |\gamma_2(j) - \gamma_2(r-j)| = O(1/rc).$$

For $\ell = 3, 4$,

$$\sum_{j=1}^{m_r}\sum_{k=1}^{m_c} [2\frac{jk}{rc}|\gamma_\ell(j,k) - \gamma_\ell(r-j, c-k)|]$$
$$= \sum_{j=1}^{m_r}\sum_{k=1}^{m_c} O(1/r)O(1/c)O(1/rc) = O(1/rc)$$

and

$$\sum_{j=1}^{m_r}\sum_{k=1}^{m_c} [2\frac{(r-j)k}{rc}|\gamma_\ell(j,k) - \gamma_\ell(j, c-k)|]$$
$$= \sum_{j=1}^{m_r}\sum_{k=1}^{m_c} (1 - O(1/r))O(1/c)O(1/rc) = O(1/c)]$$

The bound becomes $O(1/rc) + O(1/rc) + 2[O(1/r) + O(1/c) + O(1/rc)]$, which is dominated by $O(1/q)$.

Proof of Theorem 3. Without loss of generality, let us assume both r and c are odd and $c > r$. Let $m_r = [r/2]$ and $m_c = [c/2]$. The difference between Γ_1 and Γ_2 is

$$\Gamma_2 - \Gamma_1$$
$$= \sum_{k=1}^{\infty}[(\gamma(0,kc) + \gamma(kr,0))\mathrm{I}_N$$
$$+ \sum_{j=1}^{m_c}[\gamma(0,kc-j) + \gamma(0,kc+j)]\mathrm{I}_r \otimes (\Pi_c^j + \Pi_c^{jT})$$
$$+ \sum_{\ell=1}^{m_r}[\gamma(kr-\ell,0) + \gamma(kr+\ell,0)](\Pi_r^\ell + \Pi_r^{\ell T}) \otimes \mathrm{I}_c$$
$$+ \sum_{\ell=1}^{m_r}\sum_{j=1}^{m_c}[\gamma(kr+\ell,kc+j) + \gamma(kr+\ell,kc-j)$$
$$+ \quad \gamma(kr-\ell,kc+j) + \gamma(kr-\ell,kc-j)]$$
$$\times \quad (\Pi_r^\ell + \Pi_r^{\ell T}) \otimes (\Pi_c^j + \Pi_c^{jT})]$$

The largest absolute value of elements in each row corresponds to the $\gamma(kr - m_r, 0)$ or $\gamma(0, kc - m_c)$ element. It is enough to show that $\sum_{k=1}^{\infty} \gamma(kr - m_r, 0)$ goes to zero as $r \to \infty$ in order to have $|\Gamma_2 - \Gamma_1| \to 0$ element-wise. It suffices to show that the covariance function is absolutely summable. If $\sum_{j=1}^{\infty} |\gamma(j,0)| < \infty$, then

$$S_r = \sum_{k=1}^{\infty}\gamma(kr - m_r, 0) \le \sum_{k=1}^{\infty}|\gamma(kr - m_r, 0| < \infty$$

and hence $\lim_{r\to\infty} S_r = 0$.

From the expression given in (4.4) for the $CAR(1)$,

$$\sum_{k=1}^{\infty}\gamma(k,0) = \sum_{k=1}^{\infty}\sum_{s=0}^{\infty}\rho^{2s+k}\sum_{t=0}^{s}$$
$$\binom{2s+k}{2t}\binom{2t}{t}\binom{2s-2t+k}{s-t}$$

with the identity

$$\sum_{t=0}^{s}\binom{a}{t}\binom{b}{s-t} = \binom{a+b}{s},$$

we have

$$\sum_{k=1}^{\infty}\gamma(k,0) = \sum_{k=1}^{\infty}\sum_{s=0}^{\infty}\rho^{2s+k}\binom{2s+k}{s}^2.$$

With a crude upper bound on the fractorial term, $\binom{2s+k}{s} < 2^{2s+k}$, we then have

$$\sum_{k=1}^{\infty}\gamma(k,0) \le \sum_{k=1}^{\infty}\sum_{s=0}^{\infty}(4\rho)^{2s+k} = \sum_{k=1}^{\infty}(4\rho)^k\sum_{s=0}^{\infty}(16\rho^2)^s$$
$$= \frac{4\rho}{(1-4\rho)(1-16\rho^2)} < \infty.$$

Therefore, $\gamma(k,0)$ is absolutely summable

We can follow the same argument and obtain the summability result for $r > c$ and for even r or c. This implies that each element of $|\Gamma_2 - \Gamma_1| \to 0$ as $\min(r,c) \to \infty$.

ACKNOWLEDGEMENTS

This research has been supported in part by United States Department of Agriculture CSRS grant 511-100, and National Sciences Foundations grants DMS-84404970 and DMS-8704341.

BIBLIOGRAPHY

Bartlett, M.S. (1971), "Physical Nearest-Neighbour Models and Non-linear Time Series", *Journal of Applied Probability*, 8, 222-232.

Bartlett, M.S. (1978), "Nearest Neighbour Models in the Analysis of Field Experiment," *Journal of Royal Statistical Society, Series B*, 40, 147-174.

Besag, J. E. (1974), "Spatial Interaction and the Statistical Analysis of Lattice Systems," *Journal of the Royal Statistical Society, Series B*, 36, 192-236.

Besag, J. E. (1977), "Errors-in-Variables Estimation for Gaussian Lattice Schemes", *Journal of the Royal Statistical Society, Series B*, 39, 73-78.

Besag, J.E. (1981), "On a System of Two Dimensional Recurrence Equations", *Journal of Royal Statistical Society Series B*, 43, 302-309.

Besag, J. E., and Kempton, R., P.A.P. (1986), "Statistical Analysis of Field Experiments Using Neighbouring Plots", *Biometrics*, 42, 231-251.

Besag, J. E., and Moran, P.A.P. (1975), "On the Estimation and Testing of Spatial Interaction in Gaussian Lattice Processes," *Biometrika*, 62, 555-562.

Cliff, A. D., and Ord, J. K. (1981), *Spatial Processes Models and Applications*, London: Pion.

Conte, S. D., and deBoor, C. (1980), *Elementary Numerical Analysis: an Algorithmic Approach*, New York: McGraw-Hill Book Company.

Davis, P. J. (1979), *Circulant Matrices*, New York: John Wiley and Sons.

Fuller, W. (1976), *Introduction to Statistical Time Series*, New York: John Wiley and Sons.

Griffith, D. A. (1983), "The boundary value problem in spatial statistical analysis", *Journal of Region Science*, 23, 70-75.

Griffith, D.A. (1985), "An evaluation of correction techniques for boundary effects in spatial statistical analysis: contemporary methods," Geographical Analysis, 17, 1, 81-88.

Guyon, X. (1982), "Parameter Estimation for a Stationary Process on a D-Dimensional Lattice," *Biometrika*, 69(1), 95-105.

Mardia, K.V., and Marshall, R.J. (1984), "Maximum Likelihood Estimation of Models for Residual Covariance in Spatial Regression", *Biometrika*, 71(1), 135-146.

Martin, R. J. (1986), "A note on the asymptotic eigenvalues and eigenvectors of the dispersion matrix of a second-order stationary process on a *d*-dimensional lattice", *Journal of Applied Probability*, 23, 529-535.

Martin, R.J. (1987), "Some comments on correction techniques for boundary effects and missing value techniques," *Geographical Analysis*, 19, 3, 273-282.

Moran, P.A.P. (1973a), "A Gaussian Markovian Process on a Square Lattice", *Journal of Applied Probability*, 10, 54-62.

Moran, P.A.P. (1973b), "Necessary Conditions for Markovian Processes on a Lattice", *Journal of Applied Probability*, 10, 605-512.

Ord, J. K. (1975), "Estimation Methods for Models of Spatial Interaction", *Journal of American Statistical Association*, 70, 120-126.

Ripley, B. D. (1981), *Spatial Statistics*, New York: John Wiley & Sons.

Taam, W. and Yandell, B. S, (1987) "Approximate diagonalization of spatial covariance", Department of Statistics, University of Wisconsin technical report #814.

Tjostheim, D. (1978), "Statistical Spatial Series Modelling", *Advances in Applied Probability*, 10, 130-154.

Tjostheim, D. (1983), "Statistical Spatial Series Modelling II: Some Further Results on Unilateral Lattice Processes", *Advances in Applied Probability*, 15, 562-584.

Tjostheim, D. (1985), "Spatial series and time series similarities and differences", Proceedings of the 6^{th} Franco-Belgian meeting of statisticians, 217-228.

Whittle, , P. (1954), "On Stationary Processes in the Plane", *Biometrika*, 41, 434-449.