*PAPERS IN REGIONAL SCIENCE: The Journal of the RSAI* 71, 1: 71-86


# SIMPLIFYING THE NORMALIZING FACTOR IN SPATIAL AUTOREGRESSIONS FOR IRREGULAR LATTICES

**Daniel A. Griffith**
*Department of Geography*
*343 H. B. Crouse*
*Syracuse University*
*Syracuse, NY 13244-1160*
*USA*

**ABSTRACT** The Jacobian term appears in certain likelihood functions as a normalizing factor; it ensures that the use of variable transformations still leads to probability density functions whose complete integration yields unity. This term is particularly troublesome when dealing with spatial autoregressive models in that it requires numerically intensive solutions to accompanying parameter estimation problems. For these types of autoregressive models, the Jacobian term is a function of the eigenvalues of the $n$-by $n$ connectivity matrix that depicts the geographic configuration of the areal units under study. This paper reports on Jacobian approximation results, based upon supercomputer and other experiments, for irregular lattices.

## 1. INTRODUCTION

Georeferenced data, used in much geography and regional science research, introduce special complications arising from spatial dependencies that are latent in these data, and that are avoided in classical statistical analysis by invoking the assumption of independent observations. These complications cause proper spatial data analysis to be numerically intensive, and affect the quality of statistical inference decisions. One of these complications involves inclusion of a normalizing factor (called a Jacobian term) for transforming spatially autocorrelated data in such a way that reference can be made to traditional probability density functions. Ripley (1990a) argues that this particular type of numerical complexity is a major reason why researchers have shied away from spatial autoregressive modeling. Another complication involves the inclusion of spatially lagged variables, which can be computed by matrix multiplication of $n$-by-$n$ matrices times $n$-by-1 vectors, and the need to use nonlinear optimization techniques.

Anselin and Griffith (1988) have documented that the spatial statistics and spatial econometrics technology developed to date to allow proper analysis of georeferenced data are not being widely adopted. The lack of computer software to implement spatial autoregressive models appears to be a major barrier to dissemination of this technology (see Ripley, 1990b, p. 56). The principal objective of this paper is to describe simple tools that will allow spatial scientists to employ standard commercial statistical packages (such as SAS) to perform spatial

An earlier version of this paper was presented to the 37th annual North American Meeting of the Regional Science Association, Boston, November 9–11, 1990.

autoregressive modeling. The utility of these tools, which have been developed with the help of supercomputer experimentation, will be demonstrated through a comparative reanalysis of selected published regional science data sets.

## 2. BACKGROUND

Two types of spatial autoregressive models are commonly employed in the spatial statistics and spatial econometrics literature: the conditional autoregressive model (CAR), and the simultaneous autoregressive model (SAR). The major difference between these two model types stems from how their inverse covariance function is specified. Suppose the geographic configuration of areal units is depicted by the $n$-by-$n$ binary matrix $\mathbf{C}$, where entry $c_{ij} = 1$ if areal units $i$ and $j$ are juxtaposed, and $c_{ij} = 0$ otherwise. Further, suppose that matrix $\mathbf{W}$ is the stochastic version of matrix $\mathbf{C}$ (i.e., $w_{ij} = c_{ij}/\Sigma_{j=1}^{n} c_{ij}$, and thus $\Sigma_{j=1}^{n} w_{ij} = 1$). Then the inverse covariance matrix for a CAR model may be written as $(\mathbf{I} - \rho\mathbf{C})/\sigma^2$, and the inverse covariance matrix for an SAR model may be written as $(\mathbf{I} - \rho\mathbf{W})'(\mathbf{I} - \rho\mathbf{W})/\sigma^2$. One appealing feature of the CAR model is that it posits a constant conditional variance across a geographic landscape. One appealing feature of the SAR model is that it casts spatial dependence so that the value in areal unit $i$ is a weighted average of the values in juxtaposed areal units. The CAR model is appropriate for situations involving first-order dependency; the SAR model is appropriate for situations involving second-order dependency (which is captured by $\mathbf{W}'\mathbf{W}$). The focus of this paper is on the SAR model, which is more commonly used in geography and regional science; occasional references are also made to CAR model counterparts. Where application of less restrictive definitions of matrices $\mathbf{C}$ and $\mathbf{W}$ is desired (as shown in one of the ensuing reanalyzed examples), the procedures outlined in this paper should apply equally well.

The required normalizing factor is a function of the eigenvalues of matrix $\mathbf{C}$ for a CAR model, or matrix $\mathbf{W}$ for an SAR model; it may be written as follows (see Ord 1975):

$$-J_c = \sum_{i=1}^{n} \ln(1 - \rho\lambda_i)/n, \quad -J_w = 2 \sum_{i=1}^{n} \ln(1 - \rho\lambda_i)/n,$$

where $\lambda_i$ are the aforementioned eigenvalues, ln is the natural logarithm, $n$ is the number of areal units, and $\rho$ is the spatial autocorrelation parameter. Regular lattice partitionings of surfaces are characterized by systematic behavior, which is reflected in these eigenvalues. Griffith (1990a, b) has shown that these Jacobian terms can be simplified dramatically for a $\sqrt{n}$ by-$\sqrt{n}$ square tessellation, resulting in numerically simpler spatial statistical computations. This paper will explore the feasibility of Jacobian simplifications for the more common case of irregular lattice surface partitionings.

### *Findings for Regular Lattice Situations*

Griffith (1990a) has explored the constant-mean CAR model for a square-regular-lattice situation. In doing so, he has identified a function that represents the Jacobian term based on a connectivity matrix $\mathbf{C}$ for any such lattice. His

principal finding was that this Jacobian term may be quite accurately approximated by the following function:

$$J = 2\beta_n \ln(\gamma_n) - \beta_n \ln(\gamma_n + \rho) - \beta_n \ln(\gamma_n - \rho). \quad (1)$$

Selected values of $\beta_n$ and $\gamma_n$ for a square regular lattice have been estimated, and are tabulated and reported in his paper. Employment of this approximation reduces the estimation problem to one of solving a quadratic equation when only $\rho$ is unknown (or when both $\sigma$ and $\rho$ are unknown), and to one of solving a quartic equation when $\mu$ and $\rho$ are unknown (or when $\mu$, $\sigma$, and $\rho$ are unknown). Thus, because computing eigenvalues for an $n$-by-$n$ matrix is avoided, the spatial statistical problem is considerably simpler.

Next, Griffith (1990b) explored the constant-mean SAR model for a square-regular-lattice situation. Once again the functional form representing the Jacobian term was found to be that characterized by Equation (1). It may be written as

$$J = 2\alpha_n \ln(\delta_n) - \alpha_n \ln(\delta_n + \rho) - \alpha_n \ln(\delta_n - \rho). \quad (2)$$

This time, however, the stochastic version of matrix **C** · matrix **W** — was the source of eigenvalues. Estimates of $\alpha_n$ and $\delta_n$ are tabulated and reported in that paper. They exhibit the same level of accuracy previously found for the CAR model Jacobian approximation. Using this approximation reduces the estimation problem to the most common situation: that of solving a cubic equation when $\mu$, $\sigma$, and $\rho$ are unknown. Not unexpectedly, supercomputer experimental results show that as $n \to \infty$, $\alpha_n \to \beta_n$ and $\delta_n \to 4\gamma_n$.

Results for the CAR model can be generalized to a rectangular lattice through analysis of the analytical eigenvalues of matrix **C**. These eigenvalues may be written, for a $p$-by-$q$ regular lattice, as (see Ord 1975)

$$\lambda_{hk} = 2\{\cos[h\pi/(p+1)] + \cos[k\pi/(q+1)]\}, \quad h = 1, 2, \ldots, p \text{ and } k = 1, 2, \ldots, q.$$

Comparable analytical results are not yet available for matrix **W**, except for a linear geographic landscape (see Griffith 1980) where

matrix **C**: $\lambda_k = 2\{\cos[k\pi/(n+1)]\}$, $k = 1, 2, \ldots, n$, becomes

matrix **W**: $\lambda_k = \cos[k\pi/(n-1)]$, $k = 0, 1, 2, \ldots, n-1$.

Although a straightforward generalization for the SAR model is unavailable, various properties of the eigenvalues of this matrix are known. Among the more useful properties are

$P1$: $\lambda_{max} = 1$,
$P2$: $|\lambda_{min}| \leq 1$,
$P3$: $\Sigma_{i=1}^{n} \lambda_i = 0$,
$P4$: $\Sigma_{i=1}^{n} \lambda_i^2 \leq \Sigma_{i=1}^{n} 1/(\Sigma_{i=1}^{n} c_{ij})$, and
$P5$: $\lambda_{min} \leq \mathbf{u}'\mathbf{W}\mathbf{u}$ for any choice of vector $\mathbf{u}$ where $\mathbf{u}'\mathbf{u} = 1$.

These properties are useful to know because they aid in the numerical analysis

of spatial autoregression estimation. For example, in order to maintain the mathematical properties of invertibility and series convergence, $1/\lambda_{min} < \hat{\rho} < 1/\lambda_{max}$, which translates into $1/\lambda_{min} < \hat{\rho} < 1$ (using *P*1). In addition, for a regular lattice, these parameter space boundaries can be expressed as $-1 < \hat{\rho} < 1$ (using *P*2).

*The Maximum Connectivity Landscape*

Partitioning a surface into $n$ areal units can be done in numerous ways (see Griffith 1988b for small $n$ enumerations). The maximum degree of connectivity for a geographic configuration occurs in the form of an irregular lattice. The form of this lattice may be described as follows:

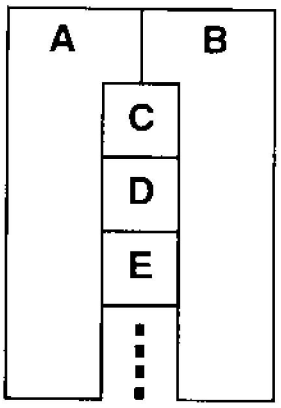
> Areal units *A* and *B* are in juxtaposition with each other as well as with all other areal units, and the remaining $n - 2$ areal units *C*, *D*, . . . form a linear arrangement having areal unit *A* adjacent to them on one side and areal unit *B* adjacent to them on the other side (see Figure 1).

Therefore, the number of connections for each areal unit is

| | |
|---|---|
| Areal unit *A* | $n - 1$ |
| Areal unit *B* | $n - 1$ |
| *C*, $n^{th}$ areal units | 3 |
| Remaining $n$-4 areal units | 4 |
| Total | $6(n - 2)$ |

The maximum eigenvalue, $\lambda_{max}$, is contained in the interval $[3, n - 1]$, for $n \geq 4$. No other geographic configuration produces a larger $\lambda_{max}$.

Because the number of connections in this maximum connectivity case is $6(n - 2)$, and the total number of possible off-diagonal entries in matrix **C** is $n(n - 1)$, as $n \to \infty$, the percentage of ones in this matrix approaches 0. This convergence suggests that $\lambda_{max}$ might converge in the limit. Both $\lambda_{max}$ and $\lambda_{min}$ have been computed for the maximum connectivity landscape, for selected values of $n$, in order to determine whether or not they converge in the limit.




| | A | B | C | D | E | . . . | ? |
|---|---|---|---|---|---|---|---|
| A | 0 | 1 | 1 | 1 | 1 | . . . | 1 |
| B | 1 | 0 | 1 | 1 | 1 | . . . | 1 |
| C | 1 | 1 | 0 | 1 | 0 | . . . | 0 |
| D | 1 | 1 | 1 | 0 | 1 | . . . | 0 |
| E | 1 | 1 | 0 | 1 | 0 | . . . | 0 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋱ | ⋮ |
| ? | 1 | 1 | 0 | 0 | 0 | . . . | 0 |

FIGURE 1. Representative Depiction of a Maximum Connectivity Partitioning of a Planar Surface

The trajectories followed by these two extreme eigenvalues appear in Figures 2a and 2b. Equations describing these trajectories were found to be the following:

$$\hat{\lambda}_{min} = 1.68446 - 1.41826\sqrt{n},$$
$$F = 1.093 \times 10^7,\ R^2 = 100\%,\ \hat{\sigma}_\varepsilon = 0.04247;\ \text{and}$$

$$\hat{\lambda}_{max} = 1.16957 + 1.42240\sqrt{n},$$
$$F = 668796.2,\ R^2 = 100\%,\ \hat{\sigma}_\varepsilon = 0.1722.$$

Clearly neither of these equations converges in the limit, meaning that the interval $1/\lambda_{min} < \rho < 1/\lambda_{max}$ shrinks to (0,0). This finding further supports use of an SAR model over a CAR model for irregular lattices. It also reflects Ord's (1975) contention that the stochastic version of matrix **C**, depicted by matrix **W**, leads to a more natural interpretation of the spatial autocorrelation parameter $\rho$.

By construction, matrix **W** is known to have $\lambda_{max} = 1$; aforementioned eigenvalue properties *P*2 and *P*3, $-1 \leq \lambda_{min} < 0$. For this maximum connectivity landscape, $\lambda_{min}$ almost immediately converges upon $-0.5$ (see Table 1). Thus, the parameter space interval for the SAR model rapidly converges upon $-2 < \rho < 1$. In contrast to the regular lattice arrangement of areal units, the irregularity of the geometric arrangement of areal units here results in a set of eigenvalues that is asymmetric about 0. Fortunately, impacts of this asymmetry on the Jacobian term quickly disappear as $n$ increases. This contention is illustrated in Figures 3a and 3b, which exhibit, respectively, the Jacobian (i.e., $-J/2$) plots for $\sqrt{n} = 2$ and $\sqrt{n} = 74$. For the irregular lattice case, then, the counterpart of Equation (2) should be of the form

$$J = \alpha_{n,1}\ln(\delta_{n,1}) + \alpha_{n,2}\ln(\delta_{n,2}) - \alpha_{n,1}\ln(\delta_{n,1} + \rho) - \alpha_{n,2}\ln(\delta_{n,2} - \rho). \tag{3}$$

This is the Jacobian approximation that is explored in this paper.

## 3. MODELS USED IN GEOGRAPHY AND REGIONAL SCIENCE

Three popular spatial autoregressive model specifications can be found in the geography and regional science literature; a complete discussion of these models is given by Upton and Fingleton (1985). The first, used for pure spatial autoregression, is the constant mean SAR model. It may be written, for some $n$-by-1 data vector **Y**, as

$$(\mathbf{Y} - \mu\mathbf{1}) = \rho\mathbf{W}(\mathbf{Y} - \mu\mathbf{1}) + \xi, \tag{4}$$

where **1** is an $n$-by-1 vector of ones, $\mu$ is the population mean, and $\xi$ is an $n$-by-1 vector of "error" values that conform to classical statistical assumptions (i.e., that they are independently and identically normally distributed, with zero mean and constant variance). This is a "constant" mean model because the population mean $\mu$ is the same for all areal units.

The second spatial autoregressive model is the variable-mean SAR model. It may be written, by attaching some set of $n$-by-$p$ predictor variable values to the vector **1** and denoting this $n$-by-$(p + 1)$ matrix with **X**, as

$$(\mathbf{Y} - \mathbf{X}\beta) = \rho\mathbf{W}(\mathbf{Y} - \mathbf{X}\beta) + \xi, \tag{5}$$

where vector $\beta$ is $(p + 1)$-by-1. Equation (5) is a multiple regression formulation

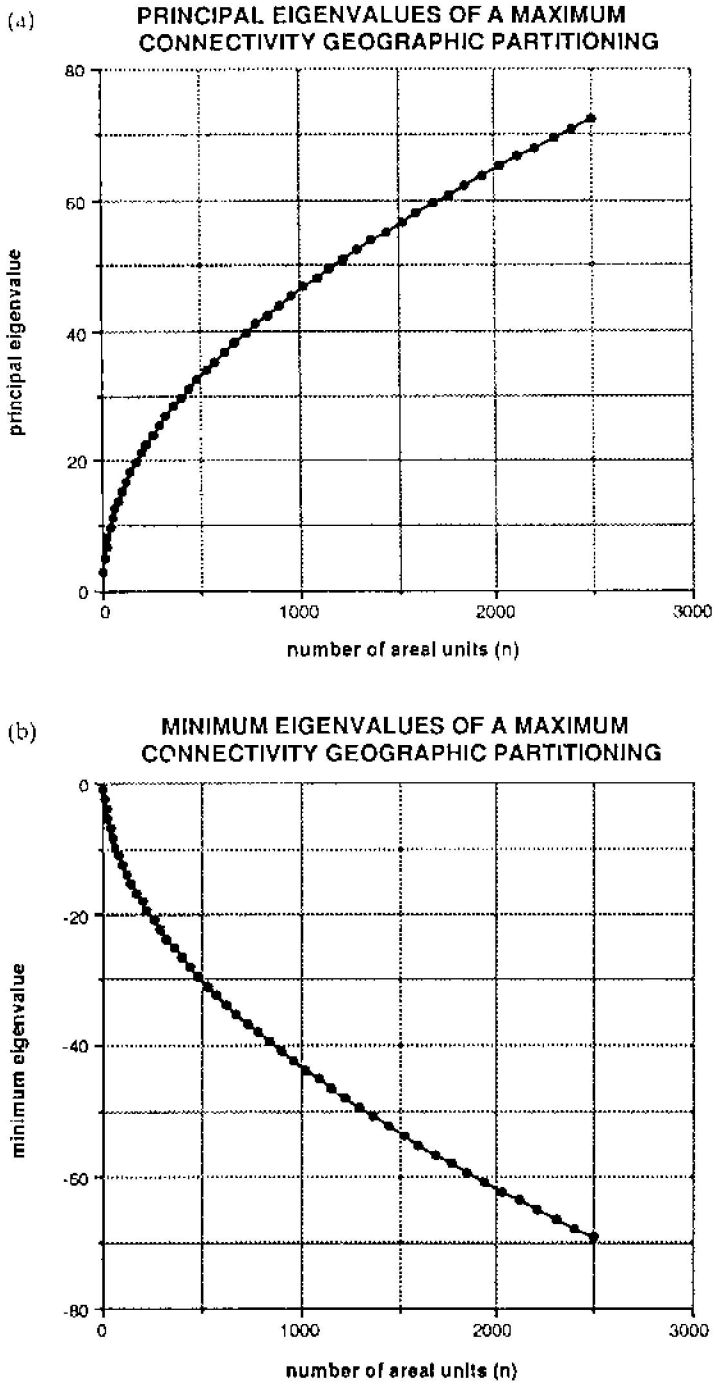



FIGURE 2. Graphical Plot of Maximum Eigenvalue (a), and Minimum Eigenvalue (b) Numerically Calculated with a Cray 2 Supercomputer from the Connectivity Matrix **C** for a Maximally Connected Surface Partitioning (See Figure 1)

TABLE 1. The Minimum Eigenvalue for the Maximum-connectivity Geographic Landscape

| $n$ | $\lambda_{min}$ | $n$ | $\lambda_{min}$ | $n$ | $\lambda_{min}$ |
|---|---|---|---|---|---|
| 4 | -0.333333333333 | 729 | −0.499999999999 | 2,601 | −0.499999999998 |
| 9 | −0.500000000000 | 784 | -0.499997676630 | 2,704 | −0.499999805275 |
| 16 | −0.493163839133 | 841 | −0.499999999999 | 2,809 | −0.499999999998 |
| 25 | −0.500000000000 | 900 | −0.499998237921 | 2,916 | −0.499999832575 |
| 36 | −0.498795569789 | 961 | −0.499999999999 | 3,025 | −0.499999999998 |
| 49 | −0.500000000000 | 1,024 | −0.499998639451 | 3,136 | −0.499999855253 |
| 64 | −0.499634116311 | 1,089 | −0.499999999999 | 3,249 | −0.499999999998 |
| 81 | −0.500000000000 | 1,156 | −0.499998932825 | 3,364 | −0.499999874217 |
| 100 | −0.499852939924 | 1,225 | −0.499999999999 | 3,481 | −0.499999999997 |
| 121 | −0.500000000000 | 1,296 | −0.499999151201 | 3,600 | −0.499999890175 |
| 144 | −0.499929804779 | 1,369 | −0.499999999999 | 3,721 | −0.499999999997 |
| 169 | −0.500000000000 | 1,444 | −0.499999316458 | 3,844 | −0.499999903680 |
| 196 | −0.499962344446 | 1,521 | -0.499999999999 | 3,969 | −0.499999999997 |
| 225 | −0.500000000000 | 1,600 | −0.499999443377 | 4,096 | −0.499999915170 |
| 256 | −0.499978015618 | 1,681 | −0.499999999999 | 4,225 | −0.499999999996 |
| 289 | −0.500000000000 | 1,764 | −0.499999542154 | 4,356 | −0.499999924999 |
| 324 | −0.499986313060 | 1,849 | −0.499999999998 | 4,489 | −0.499999999996 |
| 361 | −0.500000000000 | 1,936 | −0.499999619957 | 4,624 | −0.499999933442 |
| 400 | −0.499991037697 | 2,025 | −0.499999999998 | 4,761 | −0.499999999996 |
| 441 | −0.500000000000 | 2,116 | −0.499999681911 | 4,900 | −0.499999940732 |
| 484 | −0.499993887555 | 2,209 | −0.499999999998 | 5,041 | −0.499999999996 |
| 529 | −0.500000000000 | 2,304 | −0.499999731738 | 5,184 | −0.499999947048 |
| 576 | −0.499995688995 | 2,401 | −0.499999999998 | 5,329 | −0.499999999996 |
| 625 | −0.500000000000 | 2,500 | −0.499999772178 | 5,476 | −0.499999952545 |
| 676 | −0.499996872805 | | | | |

for which the mean for each areal unit is determined by the regression term $\mathbf{X}\beta$, a term allowing this mean to vary across a geographic landscape. For the special case of $p = 0$, $\mathbf{X} = \mathbf{1}$, and Equation (5) reduces to Equation (4). The set of $p$ predictor variables can consist of powers of areal unit centroid Cartesian coordinates, rendering a trend surface type of model. This set can consist of regional indicator variables, rendering a spatially adjusted analysis-of-variance model (see Griffith 1978). Further, this set can consist of various socioeconomic variables that covary with variate $Y$, which is the common case in regional science studies.

The third model is the spatial-autoregressive-response model. It may be written, using a multiple regression format, as

$$\mathbf{Y} = \rho\mathbf{WY} + \mathbf{X}\beta + \xi. \tag{6}$$

This model differs only slightly from, and diagnostically, is virtually indistinguishable from Equation (5). The differences between these two model specifications are, first, Equation (6) posits a more parsimonious dependence of value $y_i$ on its juxtaposed values, with the nature and degree of latent spatial autocorrelation being characterized solely by the structure of matrix $\mathbf{W}$ and the autocorrelation parameter $\rho$; and second, variable $\mathbf{Y}$ is viewed, in the usual ordinary regression model context, as a response to variables $\mathbf{X}$.

Test-statistic models are based on asymptotic variances of the various parameter estimates of Equations (4), (5), and (6). These variances, which may be computed from expectations of second partial derivatives of parameter estimates, hold for large $n$. Computation of these variances involves (a) calculating the eigenvalues of matrices $\mathbf{C}$ or $\mathbf{W}$, whichever spatial weights matrix is being used; (b) calculating, e.g., $\mathbf{WY}$; and (c) calculating an inverse matrix. All three

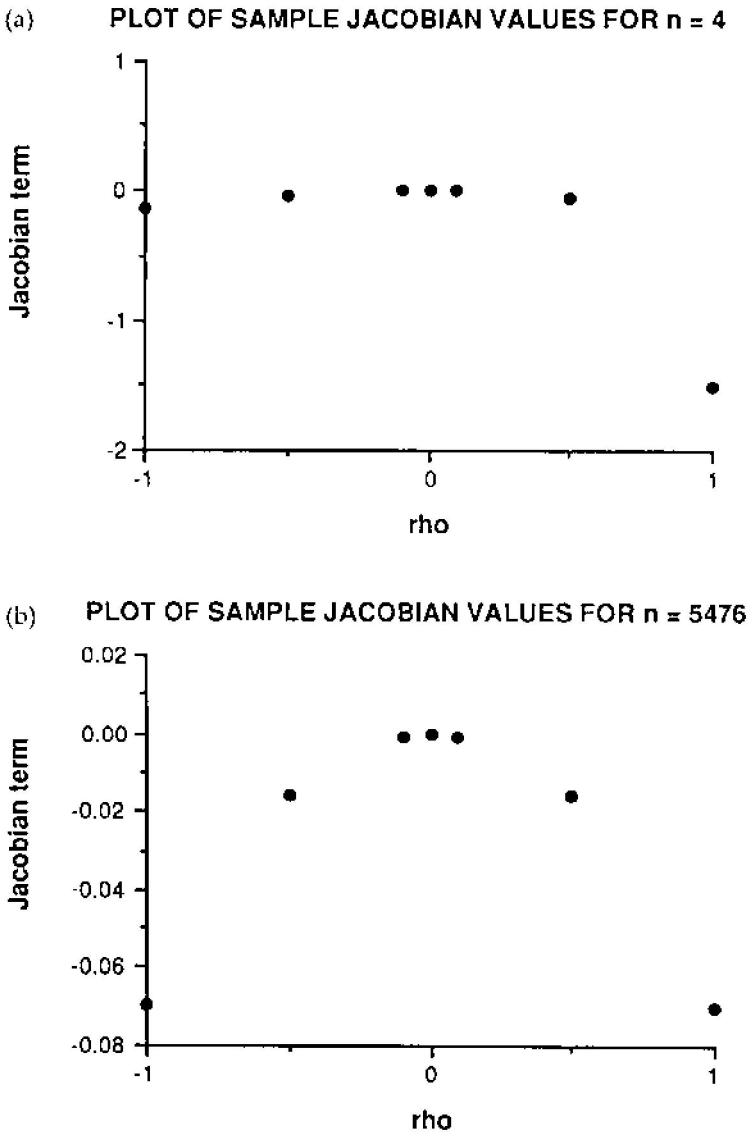



FIGURE 3. Plot of the Numerically Evaluated Jacobian Term for Selected Values of $\rho$, with (a) $\sqrt{n} = 2$, and (b) $\sqrt{n} = 74$

of these calculations cause the computing of variance estimates to be numerically intensive also.

## 4. APPROXIMATE ESTIMATION OF SPATIAL AUTOREGRESSIVE MODELS FOR IRREGULAR LATTICES

Equation (3) has been found to accurately summarize the Jacobian term for irregular lattices. This contention is demonstrated by analysis of three examples from Cliff and Ord (1981, pp. 207, 218, 219), two examples from Anselin (1988,

pp. 188, 204), and one example from Griffith (1988b, p. 219). Calibration results for these different geographic landscapes are reported in Table 2. In all cases the mean squared error is quite low; however, except for the Puerto Rico case the error is not as small for the irregular lattice case as for the regular lattice case that exploits far more systematic behavior in the underlying eigenvalues. These tabulated results suggest the following conjectures about the approximation specified with Equation (3):

(a) as $\lambda_{min} \rightarrow -1$, $\alpha_{n,1} \rightarrow \alpha_{n,2}$ and $\delta_{n,1} \rightarrow \delta_{n,2}$,
(b) $\delta_{n,2} \rightarrow 1$ (i.e., $\lambda_{max}$) from above,
(c) $\delta_{n,1} \rightarrow |1/\lambda_{min}|$ from above,
(d) as $\lambda_{min} \rightarrow -1$ and/or as $n \rightarrow \infty$, Equation (3) better approximates the Jacobian term, and
(e) as $n \rightarrow \infty$, $MAX\{\lambda_{min}\} = -0.5$ (the limiting case for the maximum connection configuration), implying that for practical sample sizes $\delta_{n,1} < 2$.

The magnitudes of the calibrated values for these four parameters display remarkable consistency over the different geographic configurations and numbers of areal units. Consequently, a generalized Jacobian approximation will also be explored. Its form is

$$J = 0.2 \cdot \ln(1.75) + 0.12 \cdot \ln(1.05) - 0.2 \cdot \ln(1.75 + \rho) - 0.12 \cdot \ln(1.05 - \rho). \quad (7)$$

In contrast to Equation (3), which still requires eigenvalues to be computed for matrix **W**, Equation (7) no longer requires such calculation.

Since $\delta_{n,2}$ appears to be governed by the principal eigenvalue, whose value is 1, and $\ln(1) = 0$, apparently the Jacobian term is impacted by $\lambda_{min}$ more than by $\lambda_{max}$. This contention is somewhat counter to Ord's (1990) claim.

Next, approximation and generalized approximation estimates were computed for twelve examples. Results of these computations are reported in Table 3. (Note that the results reported in Cliff and Ord 1981, p. 239, are incorrect; the correct results are reported in Burridge 1981. The SMSA indicator variable value for Union County in Anselin 1988, p. 205, should be 1 instead of 0.) The Puerto Rico agricultural yield example involves the constant-mean SAR model (Equation

TABLE 2. Calibration Results for the Jacobian Approximation Given by Equation (5) for Selected Geographic Landscapes

| Geographic Landscape | $n$ | $\alpha_{n,1}$ | $\alpha_{n,2}$ | $\delta_{n,1}$ | $\delta_{n,2}$ | Sum of Squared Errors | $\lambda_{min}$ | $m$ |
|---|---|---|---|---|---|---|---|---|
| Columbus, Ohio | 49 | −0.189717 | −0.120491 | 1.637114 | 1.029990 | 0.003565 | −0.650967 | 26 |
| Southwestern Ohio | 25 | −0.220243 | −0.122713 | 1.752234 | 1.016903 | 0.002324 | −0.588024 | 28 |
| Eire | 26 | −0.148738 | −0.123227 | 1.264815 | 1.016706 | 0.003430 | −0.797670 | 24 |
| Ghana | 40 | −0.181492 | −0.115805 | 1.603016 | 1.026664 | 0.004694 | −0.659530 | 27 |
| Nigeria | 50 | −0.219117 | −0.115347 | 1.873486 | 1.038506 | 0.002194 | −0.568709 | 29 |
| Puerto Rico | 73 | −0.255178 | −0.166260 | 1.754497 | 1.143221 | 0.000006 | −0.785280 | 22 |
| Regular Lattice | $\infty$ | −0.155111 | −0.155111 | 1.150234 | 1.150234 | 0.000008 | −1.000000 | 13 |
| Assumed Generalization | * | −0.2 | −0.12 | 1.75 | 1.05 | * | * | * |

*Note:* $m$ denotes the systematic sample size used to calibrate Equation (5).

TABLE 3. Comparison of Parameter Estimates for Selected Data Examples Appearing in the Geography and Regional Science Literature

| Spatial Autoregressive Model | Geographic Landscape | Parameter | MLEs | | |
|---|---|---|---|---|---|
| | | | Exact | Approximate | Approximate |
| Constant Mean | Puerto Rico | | | | |
| | DFRM | $\rho$ | 0.5519 | 0.5521 | 0.5690 |
| | | $\mu$ | 8.1869 | 8.1863 | 8.1391 |
| | DCD | $\rho$ | 0.5863 | 0.5862 | 0.5978 |
| | | $\mu$ | 35.9205 | 35.9207 | 35.9082 |
| | DMLK | $\rho$ | 0.3557 | 0.3592 | 0.3812 |
| | | $\mu$ | 27.9643 | 27.9784 | 28.0670 |
| | DSGR | $\rho$ | 0.6123 | 0.6109 | 0.6238 |
| | | $\mu$ | 7.0261 | 7.0204 | 7.0747 |
| | DCOF | $\rho$ | 0.6965 | 0.6939 | 0.6999 |
| | | $\mu$ | 4,015.3555 | 4,038.2621 | 3,988.0603 |
| | DTOB | $\rho$ | 0.5385 | 0.5390 | 0.5587 |
| | | $\mu$ | 6.2619 | 6.2617 | 6.2546 |
| | DBAN | $\rho$ | 0.5440 | 0.5445 | 0.5628 |
| | | $\mu$ | 2.6720 | 2.6710 | 2.6318 |
| | DFAM | $\rho$ | 0.4943 | 0.4959 | 0.5160 |
| | | $\mu$ | 2.1254 | 2.1247 | 2.1165 |
| Variable Mean | Eire | $\rho$ | 0.78 | 0.7737 | 0.8003 |
| | | $\beta_0$ | * | 1.2920 | 1.6526 |
| | | $\beta_1$ | 0.0030 | 0.0030 | 0.0030 |
| Response | Columbus, Ohio | $\rho$ | 0.431 | 0.4408 | 0.4464 |
| | | $\beta_0$ | 45.079 | 44.5464 | 44.2377 |
| | | $\beta_1$ | −1.032 | −1.0188 | −1.0114 |
| | | $\beta_2$ | −0.266 | −0.2657 | −0.2656 |
| Response | Southwestern Ohio (1983) | $\rho$ | −0.623 | −0.6436 | −0.6624 |
| | | $\beta_0$ | 1.604 | 1.6245 | 1.6437 |
| | | $\beta_1$ | 1.015 | 1.0167 | 1.0184 |
| | | $\beta_2$ | −0.885 | −0.8802 | −0.8759 |
| | | $\beta_3$ | −0.017 | −0.0163 | −0.0163 |
| | (1981) | $\rho$ | −0.588 | −0.6076 | −0.6259 |
| | | $\beta_0$ | 1.827 | 1.8499 | 1.8705 |
| | | $\beta_1$ | −0.731 | −0.7328 | −0.7346 |
| | | $\beta_2$ | 1.123 | 1.1251 | 1.1271 |
| | | $\beta_3$ | 0.007 | 0.0060 | 0.0050 |

*Note:* The exact MLEs for Puerto Rico data have been computed using IMSL10 subroutines (see Griffith 1990c).

4). Analysis of the density of farms (DFRM), cultivated land (DCD), milk production (DMLK), sugarcane yield (DSGR), coffee production (DCOF), tobacco production (DTOB), bananas/plantains production (DBAN), and farm families (DFAM) (all using the same **W** matrix) reveals that, for all eight variables, the approximate and the generalized approximate estimates compare extremely well with the exact maximum likelihood estimates (MLEs). This example hints that results for the two approximations are not data-specific. The generalized approximation of $\hat{\rho}$ is always greater than its exact MLE counterpart. The estimate of $\hat{\mu}$ seems to be more sensitive to the use of a Jacobian approximation than does $\hat{\rho}$.

The Eire example, which regresses percentage of gross agricultural output based on accessibility via a concomitant arterial road network, involves the variable-mean SAR model (Equation 7). Again the approximate and the generalized approximate estimates compare extremely well with the exact MLEs. In this example matrix **W** has been constructed with a function relating $w_{ij}$ to

the length of common boundaries shared by areal units and the distance between areal unit centroids. This example hints that results for these two approximations are not strictly due to the use of a stochastic version of the binary matrix **C** (i.e., $w_{ij} \neq c_{ij}/\Sigma_{j=1}^{n} c_{ij}$). Once more, the generalized approximation of $\hat{\rho}$ is greater than its exact MLE counterpart.

The Columbus, Ohio crime example and the Southwestern Ohio Phillips-curve examples involve the SAR response model (Equation 6). Again the approximate and the generalized approximate estimates compare extremely well with the MLEs. Once more, the generalized approximation of $\hat{\rho}$ is always greater than its exact MLE counterpart, and the estimate of the conditional mean ($\hat{\beta}_0$) seems to be more sensitive to the use of a Jacobian approximation than does $\hat{\rho}$. This latter contention apparently is not true for other regression parameter estimates that seem to be reasonably insensitive to the use of the Jacobian approximation. The Phillips-curve examples, which are based on a single **W** matrix, corroborate the implication from the Puerto Rico example that results for these two approximations are not data specific.

The final important feature of concern here is the estimate of standard errors, which enables hypothesis testing to be carried out. Results of these estimation exercises are reported in Table 4 for three of the example data sets. As before, the generalized approximation standard errors compare favorably with the exact asymptotic standard errors. In all cases those estimates yielded by the generalized approximation are somewhat larger, in part because the true minimum value of the likelihood function has not been attained. Therefore, the generalized approximation considerably reduces the numerical intensity of solutions, allows standard commercial software packages to be used to compute spatial statistics, and introduces a small amount of error into statistical decision making only when the test statistic is located very close to a critical value. Because standard errors have increased, statistical Type I errors will be detected slightly less often than significance levels would indicate.

The above discussion has provided the mathematical basis for undertaking spatial autoregression modeling using a Jacobian term approximation. A dis-

TABLE 4. Comparison of Asymptotic Standard Error Estimates for Selected Data Examples Appearing in the Geography and Regional Science Literature

| Geographic Landscape | Parameter | Asymptotic Standard Error | |
|---|---|---|---|
| | | Exact | Generalized Approximation |
| Eire | $\rho$ | 0.15 | 0.174 |
| | $\beta_0$ | * | 4.8988 |
| | $\beta_1$ | 0.0007 | 0.0009 |
| Columbus, Ohio | $\rho$ | 0.118 | 0.155 |
| | $\beta_0$ | 7.177 | 9.513 |
| | $\beta_1$ | 0.305 | 0.366 |
| | $\beta_2$ | 0.088 | 0.094 |
| Southwestern Ohio (1983) | $\rho$ | 0.224 | 0.289 |
| | $\beta_0$ | 0.230 | 0.296 |
| | $\beta_1$ | 0.246 | 0.274 |
| | $\beta_2$ | 0.269 | 0.303 |
| | $\beta_3$ | 0.011 | 0.013 |

cussion of how to use this method in conjunction with SAS computer statistical software appears in the Appendix of this paper.

## 5. CONCLUSIONS AND IMPLICATIONS

The procedures outlined in this paper should facilitate spatial autoregression modeling. It also should help spatial statisticians to more easily explore the analytical properties of spatial statistics.

Several unanswered research questions are pressing at this time. First, although the generalized approximation appears to be very good, it needs extensive verification, and perhaps refinement. Second, while $\lambda_{min}$ was found to be very important to the estimation process for the examples reanalyzed in this paper, this importance needs to be validated for the general irregular lattice case. In addition, a method for estimating $\lambda_{min}$ without computing the full set of $n$ eigenvalues is needed; eigenvalue properties *P*3 (a mean) and *P*4 ($n$ times a variance) suggest that this may be accomplished by establishing a general distributional form of $\lambda_i$. Ord (1990) notes that finding a vector $\mathbf{u}$ for establishing $\lambda_{max}$ is somewhat intuitive (i.e., $\mathbf{1}\sqrt{n}$, which is the principal eigenvector of matrix $\mathbf{W}$); the same cannot be said for finding a vector $\mathbf{u}$ for establishing $\lambda_{min}$ (see eigenvalue property *P*5). Finally, if $\hat{\rho}$ yielded by the generalized approximation is always biased upwards, then a correction factor needs to be established for it.

## APPENDIX

Griffith (1988a) has shown how spatial autoregressive models can be written in order to use either the MINITAB or the SAS commercial software packages to estimate parameters. One problem with SAS is that it is very difficult to incorporate the Jacobian term. The approximations outlined in this paper circumvent this particular problem, as well as making the necessary analytical derivatives less messy.

The Jacobian approximation itself can be computed as follows:

Step 1: compute the eigenvalues of matrix $\mathbf{W}$,
Step 2: compute $-J_\omega = \Sigma_{i=1}^{n} \ln(1 - \rho\lambda_i)/n$ for selected values of $\rho$ [a systematic sample should be drawn from the interval $(1/\lambda_{min}, 1)$], and
Step 3: calibrate the Jacobian approximation (Equation 3).

The four partial derivatives of Equation (3) that are necessary for its nonlinear optimization are

$$\partial J/\partial \alpha_{n,1} = \ln(\delta_{n,1}) - \ln(\delta_{n,1} + \rho)$$
$$\partial J/\partial \alpha_{n,2} = \ln(\delta_{n,2}) - \ln(\delta_{n,2} + \rho)$$
$$\partial J/\partial \delta_{n,1} = \alpha_{n,1}/(\delta_{n,1}) - \alpha_{n,1}/(\delta_{n,1} + \rho)$$
$$\partial J/\partial \delta_{n,2} = \alpha_{n,2}/(\delta_{n,2}) - \alpha_{n,2}/(\delta_{n,2} + \rho).$$

The corresponding SAS code is

```
PROC NLIN METHOD=MARQUARDT;
    PARMS A1=0.22
          A2=0.12
```

```
        D1=1.85
        D2=1.1;
    BOUNDS D1 > 1.76, D2 > 1;
    MODEL JACOB = A1*LOG(D1) + A2*LOG(D2)
                  - A1*LOG(D1 + RHO)
                  - A2*LOG(D2 - RHO);
    DER.A1 = LOG (D1) - LOG(D1 + RHO);
    DER.A2 = LOG(D2) - LOG(D2 - RHO);
    DER.D1 = A1/D1 - A1/(D1 + RHO);
    DER.D2 = A2/D2 - A2/(D2 - RHO);
RUN;
```

The variable JACOB is $-J_n = \Sigma_{i=1}^{n}\ln(1 - \rho\lambda_i)/n$, and the variable RHO consists of the systematically sampled values that were selected from the interval $(1/\lambda_{min}, 1)$ to compute JACOB. The MODEL statement is Equation (3). The four DER. statements are the partial derivatives mentioned above. The nonlinear optimization procedure, NLIN, is used here. The BOUNDS statement is based upon $\lambda_{min}$ and $\lambda_{max} = 1$. Results reported in Table 2 were computed with such SAS code; this particular code is for the Eire example.

*The Spatial Autoregressive Constant Mean Model Using Equation (4)*

Once the Jacobian has been approximated, an SAR model can be estimated. The necessary partial derivatives are as follows:

$$\partial\mathbf{Y}\cdot\exp(J)/\partial\mu = (1 - \rho)\mathbf{1}\cdot\exp(J)$$
$$\partial\mathbf{Y}\cdot\exp(J)\partial\rho = [(1 - \rho)\mu\mathbf{1} + \rho\mathbf{WY} - \mathbf{Y}]\cdot\exp(J)\cdot[-\alpha_{n,1}/(\delta_{n,1} + \rho) + \alpha_{n,2}/(\delta_{n,2} - \rho)] + (-\mu\mathbf{1} + \mathbf{WY})\cdot\exp(J).$$

The corresponding SAS code is

```
PROC NLIN METHOD=MARQUARDT MAXITER=500;
    PARMS RHO = 0.5
          B0 = 0;
    BOUNDS -1.27 < RHO < 1;
    J = .2*LOG(1.75) + .12*LOG(1.05) - .2*LOG(1.75 + RHO)
        - .12*LOG(1.05 - RHO);
    ZDFRM = (DFRM - B0*X0)*EXP(J);
    MODEL ZDFRM = RHO*(-B0*X0 + LDFRM)*EXP(J);
    DER.B0 = (1 - RHO)*X0;
    DER.RHO = (- B0*X0 + LDFRM) + (DFRM - (1 - RHO)*B0*X0
              - RHO*LDFRM)*(+.2/(1.75 + RHO) - .12/(1.05
              - RHO));
RUN;
```

RHO is the spatial autocorrelation parameter and B0 is the population mean parameter. J is the generalized approximation of the Jacobian term; X0 is a vector of ones. The MODEL statement is Equation (4). The two DER. statements are the preceding partial derivatives. The BOUNDS statement is based on $\lambda_{min}$ and $\lambda_{max} = 1$. This particular code is for the first agricultural yield variable in the Puerto Rico example.

*The Spatial Autoregressive Variable Mean Model Using Equation (5)*

$$\partial \mathbf{Y} \cdot \exp(J)/\partial \beta_0 = (1 - \rho)\mathbf{1} \cdot \exp(J)$$
$$\partial \mathbf{Y} \cdot \exp(J)/\partial \beta_j = (\mathbf{X}_j - \rho \mathbf{W}\mathbf{X}_j) \cdot \exp(J),\ j = 1, 2, \ldots, p,$$
$$\partial \mathbf{Y} \cdot \exp(J)/\partial \rho = [(\mathbf{X} - \rho \mathbf{W}\mathbf{X}\beta) + \rho \mathbf{W}\mathbf{Y} - \mathbf{Y}] \cdot \exp(J) \cdot [-\alpha_{n,1}/(\delta_{n,1} + \rho) + \alpha_{n,2}/(\delta_{n,2} - \rho)] + (-\mathbf{X}\beta + \mathbf{W}\mathbf{Y}) \cdot \exp(J).$$

The corresponding SAS code is

```
PROC NLIN METHOD=MARQUARDT MAXITER=500;
    PARMS RHO=0.5
          B0 = 0
          B1 = 1;
    BOUNDS -1 < RHO < 1;
    A1 = .2;
    A2 = .12;
    D1 = 1.75;
    D2 = 1.05;
    J = A1*LOG(D1) + A2*LOG(D2) = A1*LOG(D1 + RHO)
        - A2*LOG(D2 - RHO);
    ZY = Y*EXP(J);
    MODEL ZY = ((1 - RHO)*B0 + B1*(X - RHO*XL) + RHO*YL)*EXP(J);
    DER.B0 = (1 - RHO);
    DER.B1 = (X - RHO*XL);
    DER.RHO = ((1 - RHO)*B0 + B1*(X - RHO*XL) + RHO*YL - Y)
              *(-A1/(D1 + RHO) + A2/(D2 - RHO))
              + (- B0 - B1*XL + YL);
RUN;
```

B0 is the conditional mean parameter associated with vector **1**, and B1 is the regression coefficient for the single X variable. The spatially lagged variables are XL for X and YL for Y. The MODEL statement is Equation (5). The three DER. statements are the preceding partial derivatives. This particular code is for the Eire example.

*The Spatial Autoregressive Response Model Using Equation (6)*

$$\partial \mathbf{Y} \cdot \exp(J)/\partial \beta_0 = \mathbf{1} \cdot \exp(J)$$
$$\partial \mathbf{Y} \cdot \exp(J)/\partial \beta_j = \mathbf{X}_j \cdot \exp(J),\ j = 1, 2, \ldots, p,$$
$$\partial \mathbf{Y} \cdot \exp(J)/\partial \rho = [\mathbf{X}\beta + \rho \mathbf{W}\mathbf{Y} - \mathbf{Y}] \cdot \exp(J) \cdot [-\alpha_{n,1}/(\delta_{n,1} + \rho) + \alpha_{n,2}/(\delta_{n,2} - \rho)] + \mathbf{W}\mathbf{Y} \cdot \exp(J).$$

The corresponding SAS code is

```
PROC NLIN METHOD=MARQUARDT MAXITER=500;
    PARMS RHO=0
          B0 = 68.619
          B1 = -1.597
          B2 = -0.274;
    BOUNDS -1 < RHO < 1;
```

```
J = .189717*LOG(1.637114) + .120491*LOG(1.02999)
    - .189717*LOG(1.637114 + RHO)
    - .120491*LOG(1.02999 - RHO);
ZY = CR*EXP(J);
MODEL ZY = (B0 + B1*INC + B2*H + RHO*WCR)*EXP(J);
DER.B0 = 1;
DER.B1 = INC;
DER.B2 = H;
DER.RHO = (B0 + B1*INC + B2*H + RHO*WCR - CR)
          *(-.189717/(1.637114 + RHO) + .120491/(1.02999
          - RHO)) + WCR;
RUN;
```

J is the approximate Jacobian specific to the Columbus geographic landscape. B0 is the conditional mean parameter associated with vector **1**, and B1 and B2 are the regression coefficients for the two X variables (i.e., INC and H). The spatially lagged variable for CR is WCR. The MODEL statement is Equation (6). The three DER. statements are the preceding partial derivatives. This particular code is for the Columbus, Ohio crime example.

Because both sides of the preceding sets of partial derivatives are multiplied by $\exp(J)$, this term can be excluded from the SAS DER. statements. The spatially lagged terms **WX** and **WY** are best computed as a data preprocessing task. The initial parameter guesses for PROC NLIN are crucial. Often ordinary least-squares results can be used, as has been done here with the Columbus, Ohio example; however, this strategy was unsuccessful with the Eire example discussed in this paper. Since the spatial autoregression normally would follow an ordinary least-squares solution, coupled with a test for spatially autocorrelated residuals, a reasonable initial value can be ascertained from the spatial autocorrelation diagnostic statistic (e.g., the Moran Coefficient). A convenient initial value of RHO is $\pm 0.5$. Griffith and Amrhein (1991) have furnished sample SAS code illustrating how to complete the data preprocessing task, and how to compute the Moran Coefficient for standard regression residuals.

## ACKNOWLEDGMENTS

This research was supported by NSF grant SES-87-22086. Computing was sponsored by a grant from the National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign.

## REFERENCES

Anselin, L. 1988. *Spatial econometrics: methods and models*. Dordrecht, Netherlands: Kluwer.

Anselin, L., and Griffith, D. 1988. Do spatial effects really matter in regression analysis? *Papers of the Regional Science Association* 65: 11–34.

Burridge, P. 1981. Testing for a common factor in a spatial autoregression model. *Environment and Planning A* 13: 795–800.

Cliff, A., and Ord, J. 1981. *Spatial processes*. London: Pion.

Griffith, D. 1978. A spatially adjusted ANOVA model. *Geographical Analysis* 10: 296–301.

Griffith, D. 1980. Towards a theory of spatial statistics. *Geographical Analysis* 12: 325–39.

Griffith, D. 1988a. Estimating spatial autoregressive model parameters with commercial statistical packages. *Geographical Analysis* 20: 176–86.

Griffith, D. 1988b. *Advanced spatial statistics*. Dordrecht, Netherlands: Kluwer.

Griffith, D. 1990a. A numerical simplification for estimating parameters of spatial autoregressive models. In *Spatial statistics: past, present, and future*, ed. D. Griffith, pp. 183–97. Ann Arbor, Michigan: Institute of Mathematical Geography.

Griffith, D. 1990b. Simplifying the normalizing factor in spatial autoregression. Paper delivered at the advanced workshop on statistical modeling sponsored by the International Geographical Union Commission on Mathematical Models, Boston University, Boston, Massachusetts.

Griffith, D. 1990c. Supercomputing and spatial statistics: a reconnaissance. *The Professional Geographer* 42: 481–92.

Griffith, D., and Amrhein, C. 1991. *Statistical analysis for geographers*. Englewood Cliffs, New Jersey: Prentice-Hall.

Ord, J. 1975. Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70: 120–26.

Ord, K. 1990. Disucssion. In *Spatial statistics: past, present, and future*, ed. D. Griffith, pp. 197–99. Ann Arbor, Michigan: Institute of Mathematical Geography.

Ripley, B. 1990a. Gibbsian interaction models. In *Spatial statistics: past, present, and future*, ed. D. Griffith, pp. 3–25. Ann Arbor, Michigan: Institute of Mathematical Geography.

Ripley, B. 1990b. Discussion. In *Spatial statistics: past, present, and future*, ed. D. Griffith, pp. 55–57. Ann Arbor, Michigan: Institute of Mathematical Geography.

Upton, G., and Fingleton, B. 1985. *Spatial data analysis by example*, vol. 1. New York: Wiley.