SINET: Ethiop. J. Sci., 24(1):17–33, 2001
 ISSN: 0379–2897

THE EFFICIENCY OF OLS ESTIMATOR IN THE LINEAR REGRESSION MODEL WITH SPATIALLY CORRELATED ERRORS

Butte Gotu

Department of Statistics, Faculty of Science, Addis Ababa University
PO Box 1176, Addis Ababa, Ethiopia, E-mail: Sinet-aau@telecom.net.et

ABSTRACT: Bounds for the efficiency of ordinary least squares estimator relative to generalized least squares estimator in the linear regression model with first-order spatial error process are given.

Key words/phrases: Efficiency, generalized least squares, ordinary least squares, spatial error process, spatial correlation

INTRODUCTION

Let the relationship between an observable random variable y and k explanatory variables $X_1,\ldots,X_k$ in a T-county system be specified in linear regression form

$$y = X\beta + u, \tag{1}$$

where X is a $T \times k$ matrix of known constants with full column rank $k < T$, and β is a $k \times 1$ vector of unknown parameters. The vector u is a disturbance term with $E(u) = 0$ and $Cov(u) = \sigma_\epsilon^2 V_*$, where σ_ϵ^2 is a positive unknown scalar and V_* a $T \times T$ positive definite matrix with identical diagonal elements. The assumption that the diagonal elements of V_* are all identical indicates that we consider only homoscedastic disturbances $u'_i s$ which are correlated.

The ordinary least squares (OLS) and the generalized least squares (GLS) estimators of β in model (1) are given by $\hat{\beta} = (X'X)^{-1}X'y$ and $\tilde{\beta} = (X'V_*^{-1}X)^{-1}X'V_*^{-1}y$, respectively, with covariance matrices $Cov(\hat{\beta}) = \sigma_\epsilon^2(X'X)^{-1}X'V_*X(X'X)^{-1}$ and $Cov(\tilde{\beta}) = \sigma_\epsilon^2(X'V_*^{-1}X)^{-1}$.

When the covariance of the disturbance vector u is not a scalar multiple of the identity matrix, that is $Cov(u) \neq \sigma_\epsilon^2 I$ as in model (1), it is well known that the GLS estimator provides the best linear unbiased estimator (BLUE) of β in contrast to OLS (see Fomby *et al.*, 1984, p. 17).

But in applications, Cov(u) usually involves unknown parameters like a spatial correlation coefficient, so one has to look for another estimator, OLS, say. In cases where Cov(u) does not involve unknown parameters, one problem facing a researcher dealing with model (1) is how to measure the efficiency of OLS estimator $\hat{\beta}$ relative to GLS estimator $\tilde{\beta}$. For spatial case, this question can be expressed as: what can we gain by estimating β in the regression model based on spatial assumptions instead of using simple standard regression specifications?

A number of authors have investigated the efficiency of OLS relative to GLS estimator when the errors are serially or spatially correlated by using various efficiency criteria (see Bloomfield and Watson, 1975; Krämer, 1980; Krämer and Donninger, 1987; Haining, 1990; Griffith, 1988; Cordy and Griffith, 1993; Krämer and Baltagi, 1996). The most remarkable feature of the results obtained is that the relative efficiency depends mainly on the error process considered, the regressor matrix X and the degree of correlation. The relevant literature can be grouped into two, namely papers which assume X as given and seek bounds for the efficiency as the covariance matrix varies, and papers which take the covariance matrix as given and consider bounds for the efficiency as X varies. Another aspect of the resulting analysis shows the behaviour of the relative efficiency of OLS when the correlation parameter tends toward the boundary of the parameter space.

In this paper, bounds for the efficiency of OLS relative to GLS estimator of β in model (1) under first-order spatial error process, that is under the assumption that the covariance matrix is given, are constructed by using the measures of efficiency based on

- the euclidean norm of the difference $P_X V_* - V_* P_X$, $P_X = X(X'X)^{-1}X'$
- the ratio of the traces of the covariance matrices of $X\tilde{\beta}$ and $X\hat{\beta}$
- the ratio of the determinants of the covariances of $\tilde{\beta}$ and $\hat{\beta}$.

The bounds obtained can then be used to decide whether to apply OLS method or GLS method based on the assumptions of first-order spatial process.

LINEAR REGRESSION WITH FIRST-ORDER SPATIAL ERROR PROCESSES

In order to analyze the efficiency of OLS relative to GLS estimator given specific error process, one needs the structure of the covariance matrix of the disturbance vector u. So, we start by specifying first-order spatial error processes.

Let the components of u follow a first-order spatial moving average (MA(1)) process

$$u_i = \rho \sum_{j=1}^{T} w_{ij}\epsilon_j + \epsilon_i$$

or, in matrix form

$$u = \rho W\epsilon + \epsilon \tag{2}$$

where ρ denotes a spatial correlation coefficient for a given county system and ϵ is an error term with $E(\epsilon) = 0$ and $Cov(\epsilon) = \sigma_\epsilon^2 I$ (I is the T-dimensional identity matrix). W is a $T \times T$ matrix whose elements are known non-negative weights defined by (Cliff and Ord 1981, pp. 17–19):

$$w_{ij} \begin{cases} >0, & \textit{if } i \textit{ and } j \textit{ are neighbours } (i \neq j) \\ =0, & \textit{otherwise.} \end{cases}$$

The element w_{ij} of the weights matrix W measures the strength of the effect of county j on county i.

Under first-order spatial autoregressive (AR(1)) process, the components of u follow the pattern

$$u_i = \rho \sum_{j=1}^{T} w_{ij} u_j + \epsilon_i$$

or, in matrix form

$$u = \rho W u + \epsilon . \tag{3}$$

Equations (2) and (3) can be written as

$$u = (I + \rho W)\epsilon \text{ and } u = (I - \rho W)^{-1} \epsilon , \tag{4}$$

respectively, where in the AR(1) case the matrix $I - \rho W$ must be non-singular. From (1) and (4), we obtain four possible structures of $\boldsymbol{Cov(u)} = \sigma_\epsilon^2 \boldsymbol{V}_*$ for the first-order spatial error process:

$$V_* = \begin{cases} (I+\rho W)(I+\rho W') & : \; MA(1) \\ (I+\rho W) & : \; MA(1) - conditional \\ (I-\rho W)^{-1}(I-\rho W')^{-1} & : \; AR(1) \\ (I-\rho W)^{-1} & : \; AR(1) - conditional. \end{cases} \tag{5}$$

The conditional cases are special cases of the unconditional process (Bartlett, 1971; Besag, 1974). To ensure that V_* is positive definite, the possible values of ρ must be identified (see Horn and Johnson, 1985, p. 301). Once the structure of the covariance matrix is specified the next step will be to analyze the efficiency of OLS estimator by using different efficiency measures. Before we analyze efficiency, let us introduce a general expression of the covariance matrix of the disturbance vector and give some general assumptions used throughout the paper.

According to the assumptions given in model (1) the matrix V_* has identical diagonal elements, and denoting this element by υ, we get

$$\boldsymbol{Cov(u)} = \sigma_\epsilon^2 \boldsymbol{V}_* = (\upsilon\, \sigma_\epsilon^2)\boldsymbol{V} = \sigma_u^2 \boldsymbol{V}, \tag{6}$$

where $\boldsymbol{V} = (1/\upsilon)\boldsymbol{V}_*$ and $\sigma_u^2 = \upsilon\, \sigma_\epsilon^2$ is the variance of the disturbances u_i, $i = 1, \ldots , T$. Using the above assumptions under spatial process we can now write model (1) as the familiar general linear regression model

$$y = X\beta + u, \quad E(u) = 0, \quad Cov(u) = \sigma_u^2 V. \tag{7}$$

We assume that $X'X = I$. For the matrix X with full column rank, there is no loss of generality in assuming that $X'X=I$ because under the transformation

$$y = \tilde{X}\delta + u \tag{8}$$

with $\tilde{X} = X(X'X)^{-1/2}$ and $\delta = (X'X)^{1/2}\beta$ the condition $\tilde{X}'\tilde{X} = I$ is valid for all X, and the OLS and GLS estimators of β are given by $\hat{\beta} = (X'X)^{-1/2}\,\hat{\delta}$ and $\tilde{\beta} = (X'X)^{-1/2}\,\tilde{\delta}$, respectively. $\hat{\delta}$ and $\tilde{\delta}$ are the estimators of δ in (8). Furthermore, we impose the following assumptions on the weight matrix W throughout or when needed (see Appendix regarding examples):

- W is symmetric.
- W is symmetric and orthogonal.
- W is symmetric with row sums equal to one.

EFFICIENCY BASED ON THE EUCLIDEAN NORM

Consider the measure of efficiency based on the euclidean norm of the difference $P_X V - VP_X$ defined by (see Bloomfield and Watson, 1975)

$$\begin{aligned} e_1(\rho) &:= \tfrac{1}{2}\,\| P_X V - VP_X \|^2 \\ &= \tfrac{1}{2} tr((P_X V - VP_X)'(P_X V - VP_X)) \\ &= tr(P_X V^2) - tr(P_X V)^2 . \end{aligned} \tag{9}$$

When $e_1(\rho) = 0$, the OLS estimator $\hat{\beta}$ can be applied without loss of efficiency whereas a loss of efficiency is expected if $e_1(\rho) \neq 0$. In what follows $\mu_i(A)$ denotes the i-th eigenvalue of a $T \times T$ matrix A.

Theorem 1 (Bloomfield and Watson, 1975)

Assume that $\mu_1(V) \leq \ldots \leq \mu_T(V)$. Under the assumptions that $X'X = I$, V positive definite and $T \geq 2k$, we have

$$e_1(\rho) \le \frac{1}{4}\sum_{i=1}^{k}(\mu_i(V) - \mu_{T-i+1}(V))^2 . \tag{10}$$

<u>*Remarks:*</u>
When there are big differences within the k pairs ($\mu_i(V)$, $\mu_{T-i+1}(V)$) of the eigenvalues of V, then the bound in (10) will be large. The restriction $T \geq 2k$ by Bloomfield and Watson allows to take pairs at a time.

Based on the theorem by Bloomfield and Watson the following result can be stated.

Theorem 2
Let $\mu_1(V) \leq \ldots \leq \mu_T(V)$. Under the assumptions that $X'X = I$, V positive definite and $T \geq 2k$, we obtain

$$e_1(\rho) \le \frac{1}{4\upsilon^2}\sum_{i=1}^{k}(\mu_i(V_*) - \mu_{T-i+1}(V_*))^2 , \tag{11}$$

Proof:
By inserting $V = (1/\upsilon)V_*$ in (9), we have

$$e_1(\rho) = \frac{1}{\upsilon^2}\{tr(P_X V_*^2) - tr(P_X V_*)^2\} . \tag{12}$$

The result follows then by applying Theorem 1. ◇

In the following the upper bounds of $e_1(\rho)$ will be given by applying the relationship given in (11) under some assumptions on the weights matrix.

Corollary 1
Let $X'X = I$ and $T \geq 2k$. When the components of the disturbance vector u in model (7) follow a conditional spatial $MA(1)$ process, then

$$e_1(\rho) \le \frac{\rho^2}{4}\sum_{i=1}^{k}(\mu_i(W) - \mu_{T-i+1}(W))^2 . \tag{13}$$

Proof:

For a conditional spatial moving average process of order one the matrix V_* is given by $V_* = (I + \rho W)$, with W being symmetric. The diagonal elements of V_* are all equal to one because the respective elements of the weights matrix are all equal to zero. This implies that $\upsilon = 1$. Furthermore,

$$\mu_i(V_*) = 1 + \rho\, \mu_i(W) \, , \tag{14}$$

where the eigenvalues $\mu_i(V_*)$ and $\mu_i(W)$, $i=1, \ldots , T$ are in ascending order. Inserting (14) in (11) completes the proof. ◇

The bound in (13) will be large when there are large differences within the k pairs of eigenvalues $(\mu_i(W), \mu_{T-i+1}(W))$ of the matrix W. That is, the efficiency of OLS relative to GLS estimator will be lower when the differences within the pairs of eigenvalues of W are large.

If the row sums of W are equal to one, then $e_1(\rho) \leq k\rho^2$ because the absolute value of the eigenvalue $\mu_i(W)$ is less than or equal to one for all i (see Graybill, 1983, p. 98).

Remark:

The result of Corollary 1 also holds for a conditional spatial AR(1) process if W is orthogonal.

Corollary 2

Assume that W is orthogonal and symmetric. Let $X'X = I$ and $T \geq 2k$. When the components of the disturbance vector u in model (7) follow a spatial $MA(1)$ or $AR(1)$ process, then

$$e_1(\rho) \leq \frac{4k\rho^2}{(1+\rho^2)^2} \; .$$

Proof:

***MA*(1) process:**

Under a spatial $MA(1)$ process we have

$$V_* = (I + \rho W)(I + \rho W').$$

From the assumption that the weights matrix W is orthogonal and symmetric it follows that

$$V_* = (1 + \rho^2)I + 2\rho W,$$

implying $\upsilon = 1 + \rho^2$ and $\mu_i(V_*) = (1 + \rho^2) + 2\rho\mu_i(W)$. Inserting these eigenvalues in (11) we get

$$e_1(\rho) \leq \frac{\rho^2}{(1 + \rho^2)^2} \sum_{i=1}^{k} (\mu_i(W) - \mu_{T-i+1}(W))^2$$

Since W is orthogonal and symmetric we have $\mu_i(W) \in \{-1,1\}$, which gives $e_1(\rho) \leq (4k\rho^2)/(1 + \rho^2)^2$

AR(1) process:

Under a spatial $AR(1)$ process the matrix V_* is given by

$$V_* = (I - \rho W)^{-1}(I - \rho W')^{-1}.$$

When the weights matrix W is assumed to be symmetric and orthogonal, we obtain $(I - \rho W)^{-1} = (1/(1 - \rho^2))(I + \rho W)$ (see Searle, 1982, p. 137), and V_* has the form

$$V_* = \frac{1}{(1-\rho^2)^2}((1+\rho^2)I + 2\rho W))$$

This implies that $\upsilon = (1 + \rho^2)/(1 - \rho^2)^2$ and

$$\mu_i(V_*) = \frac{1}{(1-\rho^2)^2}((1+\rho^2) + 2\rho\,\mu_i(W)) , \qquad (15)$$

where the eigenvalues are in ascending order. Inserting (15) in (11) and using the fact that $\mu_i(W) \epsilon \{-1,1\}$ completes the proof. ◇

The following result shows that the OLS estimator can be applied without loss of efficiency as ρ goes to one.

Theorem 3

Let $\Re(X)$ be the k-dimensional space spanned by the columns of X, and let $\ell:=(1,\ldots,1)' \in \Re(X)$. If $\lim_{\rho\to 1} V = c\ell\ell'$, $c \in \mathbb{R}$, then $\lim_{\rho\to 1} e_1(\rho) = 0$.

Proof:

The efficiency $e_1(\rho)$ can be written as:

$$e_1(\rho) = tr(P_X V^2) - tr(P_X V)^2 = tr(P_X V(V - P_X V))$$
$$= tr(P_X V M_X V).$$

When the condition $\lim_{\rho\to 1} V = c\ell\ell'$ holds, we have

$$\lim_{\rho \to 1} e_1(\rho) = c^2 tr(P_X \ell\ell' M_X \ell\ell')$$

Since $\ell \in \Re(X)$ we get $M_X\ell = (I - P_X)X\gamma = 0$, γ being a $k\times 1$ vector, and this implies $\lim_{\rho\to 1} e_1(\rho) = 0$. ◇

EFFICIENCY BASED ON THE RATIO OF TRACES

If the ratio of the mean squared errors is used to define the measure of efficiency of OLS relative to GLS estimator, then we have (see Krämer, 1980)

$$e_2(\rho) := \frac{tr(Cov(X\tilde{\beta}))}{tr(Cov(X\hat{\beta}))}$$

with $Cov(X\tilde{\beta}) = \sigma_u^2 X(X'V^{-1}X)^{-1}X'$ and $Cov(X\hat{\beta}) = \sigma_u^2 P_X V P_X$. Using this measure of efficiency a number of papers investigate the efficiency of OLS relative to GLS estimator under stationary AR(1) process in time series and spatial models (see Krämer, 1980, 1984; Krämer and Donninger, 1987; Krämer and Baltagi, 1996).

The following theorem gives a lower bound for $e_2(\rho)$ which holds for all covariance structures under general linear regression model (7).

Theorem 4

Let $X'X = I$. Then

$$\frac{\sum_{i=1}^{k} \mu_i(V)}{\sum_{i=1}^{k} \mu_{T-k+i}(V)} \leq e_2(\rho) \leq 1 \ . \tag{16}$$

Proof:

Since σ_u^2, in $e_2(\rho)$, cancels out, we set $\sigma_u^2 = 1$ in calculating covariances. Under the assumption $X'X = I$, we have

$$tr(Cov(X\hat{\beta})) = tr(P_X V P_X) = tr(X'VX) \tag{17}$$

and

$$\begin{aligned} tr(Cov(X\tilde{\beta})) &= tr(X(X'V^{-1}X)^{-1}X') = tr(X'V^{-1}X)^{-1} \\ &= \sum_{i=1}^{k} \mu_i((X'V^{-1}X)^{-1}) \\ &= \sum_{i=1}^{k} \frac{1}{\mu_i(X'V^{-1}X)} \ . \end{aligned} \tag{18}$$

Applying Poincaré separation theorem we obtain the following inequalities (see Horn and Johnson, 1985, p. 190):

$$\begin{aligned} &\sum_{i=1}^{k} \mu_i(V) \leq tr(Cov(X\hat{\beta})) \leq \sum_{i=1}^{k} \mu_{T-k+i}(V) \\ &\mu_i(V^{-1}) \leq \mu_i(X'V^{-1}X) \leq \mu_{T-k+i}(V^{-1}) \ . \end{aligned} \tag{19}$$

The second inequality in (19) implies

$$\frac{1}{\mu_i(X'V^{-1}X)} \geq \frac{1}{\mu_{T-k+i}(V^{-1})}, \quad i = 1,\ldots,k \ .$$

Using (17) to (19) we have

$$tr(Cov(X\tilde{\beta})) \geq \sum_{i}^{k} \frac{1}{\mu_{T-k+i}(V^{-1})}$$

$$= \sum_{i}^{k} \mu_i(V)$$

$$tr(Cov(X\hat{\beta})) \leq \sum_{i=i}^{k} \mu_{T-k+i}(V) \ . \qquad (20)$$

From (20) it is clear that

$$\frac{\sum_{i=1}^{k} \mu_i(V)}{\sum_{i=1}^{k} \mu_{T-k+i}(V)} \leq e_2(\rho) \ .$$

The inequality $e_2(\rho) \leq 1$ follows from the optimality of GLS estimator (see Krämer, 1980). ◇

If there is a large difference between the sum of the k smallest and k largest eigenvalues of V, then the efficiency of OLSE will significantly be smaller, but never less than the ratio of the smallest and the largest eigenvalues $\mu_{min}(V)/\mu_{max}(V)$.

Remark:
If the diagonal elements of V_* are not identical, meaning that the $u'_i s$ have different variances, then we get

$$e_2(\rho) \geq \frac{\sum_{i=1}^{k} \mu_i(V_*)}{\sum_{i=1}^{k} \mu_{T-k+i}(V_*)} \ .$$

For spatial models with first-order spatial error process the following result is obtained.

Corollary 3

Assume that the matrix X fulfils $X'X = I$. Let the weights matrix W be symmetric with row sums equal to one. If the components of the disturbance vector u follow a spatial $MA(1)$ or $AR(1)$ process, then

$$e_2(\rho) \geq \frac{(1-\rho)^2}{(1+\rho)^2}, \qquad \rho > 0 . \tag{21}$$

Proof:

MA(1) process

Under a spatial $MA(1)$ process with symmetric weights matrix the eigenvalues of V_* are given by

$$\mu_i(V_*) = (1 + \rho\, \mu_i(W))^2 , \qquad i = 1,\ldots,T ,$$

where the eigenvalues of W and V_* are in ascending order. When the row sums of W are all equal to one, then the absolute value of $\mu_i(W)$ is less than or equal to one for all i (see Graybill, 1983, p. 98). This implies

$$\frac{1}{\upsilon}(1-\rho)^2 \leq \mu_i(V) \leq \frac{1}{\upsilon}(1+\rho)^2 , \qquad \rho > 0 , \tag{22}$$

so that applying Theorem 4 gives (21).

AR(1) process

Using the same reasoning as in the $MA(1)$ case we obtain the following bounds for the eigenvalues of V:

$$\frac{1}{\upsilon(1+\rho)^2} \leq \mu_i(V) \leq \frac{1}{\upsilon(1-\rho)^2} , \qquad \rho > 0 \tag{23}$$

and (21) follows by applying Theorem 4.

In what follows we use a measure of efficiency which is based on the determinants of the covariances of the least squares estimators, and give a lower bound for the efficiency of OLS relative to GLS estimator.

EFFICIENCY BASED ON THE RATIO OF DETERMINANTS

Consider the measure of efficiency given by (see Watson, 1955)

$$e_3(\rho) := \frac{|Cov(\tilde{\beta})|}{|Cov(\hat{\beta})|} = \frac{|X'X|^2}{|X'VX|\,|X'V^{-1}X|},$$

where $|.|$ stands for determinant. The matrices $X'VX$ and $X'V^{-1}X$ are positive definite because V is positive definite and X of full column rank. This implies that $e_3(\rho) > 0$.

Let A and B be T x k matrices and assume that $B'B$ is non-singular. The well known Cauchy-Inequality concerning the determinants of two matrices A and B states that $|A'B|^2 \leq |A'A|\ |B'B|$ (see Basilevsky, 1983, p. 167). Using $A = V^{1/2}X$ and $B = V^{-1/2}X$, we get $|X'X|^2 \leq |X'VX|\,|X'V^{-1}X|$. This implies, under the assumption $X'X = I$, $e_3(\rho) \leq 1$.

The following theorem gives a lower bound for $e_3(\rho)$.

Theorem 5
Let $X'X = I$. Then

$$e_3(\rho) \geq \prod_{i=1}^{k} \frac{\mu_i(V)}{\mu_{T-k+i}(V)}. \tag{24}$$

Proof:
By applying Poincaré separation theorem we get

$$\prod_{i=1}^{k} \mu_i(V) \le \prod_{i=1}^{k} \mu_i(X'VX) \le \prod_{i=1}^{k} \mu_{T-k+i}(V)$$
$$\prod_{i=1}^{k} \mu_i(V^{-1}) \le \prod_{i=1}^{k} \mu_i(X'V^{-1}X) \le \prod_{i=1}^{k} \mu_{T-k+i}(V^{-1}) ,$$

where the eigenvalues are in ascending order. This implies

$$|X'VX| = \prod_{i=1}^{k} \mu_i(X'VX) \ge \prod_{i=1}^{k} \mu_i(V) ,$$

so that

$$\frac{1}{|X'VX|} \le \prod_{i=1}^{k} \frac{1}{\mu_i(V)} .$$

Furthermore,

$$|X'VX| \le \prod_{i=1}^{k} \mu_{T-k+i}(V)$$
$$|X'V^{-1}X| \le \prod_{i=1}^{k} \mu_{T-k+i}(V^{-1}) .$$

This implies

$$\frac{1}{|X'VX|} \ge \prod_{i=1}^{k} \frac{1}{\mu_{T-k+i}(V)}$$
$$|X'V^{-1}X| \le \prod_{i=1}^{k} \frac{1}{\mu_i(V)} . \qquad (25)$$

According to the definition, we have

$$e_3(\rho) = \frac{1/|X'VX|}{|X'V^{-1}X|} ,$$

and using (25) yields the asserted result. ◇

Remark:
Bloomfield and Watson (1975) give a narrower lower bound for $e_3(\rho)$ under the additional assumptions that $T \geq 2k$ and $k > 1$.

Under first-order spatial error process we get the following result.

Corollary 4
Assume that $X'X = I$. Let the weights matrix W be symmetric with row sums equal to one. If the components of the disturbance vector u follow a spatial $MA(1)$ or $AR(1)$ process, then

$$e_3(\rho) \geq \frac{(1-\rho)^{2k}}{(1+\rho)^{2k}}, \quad \rho > 0 .$$

Proof:
The proof follows by applying Theorem 5 using the bounds of the eigenvalues of the matrix V given in (22) and (23). ◇

Remark:
When the diagonal elements of V_* are not identical, meaning that the u_i's have different variances, we get

$$e_3(\rho) \geq \prod_{i=1}^{k} \frac{\mu_i(V_*)}{\mu_{T-k+i}(V_*)} .$$

REFERENCES

1. Bartlett, M.S. (1971). Physical nearest neighbour models and non-linear time series. *Journal of Applied Probability* **8:**222–232.
2. Besag, J.E. (1974). Spatial interaction and statistical analysis of lattice systems, *Journal of Royal Statistical Society* B **36:**192–236.
3. Basilevsky, A. (1983). *Applied Matrix Algebra in the Statistical Sciences.* Elsevier Science Publishing, New York.
4. Bloomfield, P. and Watson, G.S. (1975). The inefficiency of least squares. *Biometrika* **62:**121–128.
5. Cliff, A.D. and Ord, J.K. (1981). *Spatial Processes: Models and Applications.* Pion, London.

6. Cordy, C.B. and Griffith, D.A. (1993). Efficiency of least squares estimators in the presence of spatial autocorrelation. *Communications in Statistics, Simulation and Computation* **22**:1161–1179.

7. Fomby, T.B., Hill, R.C. and Johnson, S.R. (1984). *Advanced Econometric Methods*. Springer-Verlag, New York.

8. Graybill, F.A. (1983). *Matrices with Applications in Statistics*. 2nd ed., Wadsworth, Belmont.

9. Griffith, D.A. (1988). *Advanced Spatial Statistics*. Kluwer, Dordrecht.

10. Haining, R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge.

11. Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge University Press, Cambridge.

12. Krämer, W. (1980). Finite sample efficiency of ordinary least squares in the linear regression model with autocorrelated errors. *Journal of the American Statistical Association* **75**:1005–1009.

13. Krämer, W. (1984). High correlation among Errors and the efficiency of ordinary least squares in the linear models. *Statistical Papers* **25**:135–142.

14. Krämer, W. and Donninger, C. (1987). Spatial autocorrelation among errors and the relative efficiency of OLS in the linear regression model. *Journal of the American Statistical Association* **82**:577–579.

15. Krämer, W. and Baltagi, B. (1996). A general condition for an optimal limiting efficiency of OLS in the general linear regression model. *Economic Letters* **50**:13–17.

16. Searle, S.R. (1982). *Matrix Algebra Useful for Statistics*. Wiley, New York.

17. Watson, G.S. (1955). Serial correlation in regression analysis. *Biometrika* **42**:327–341.

Appendix

Examples for orthogonal and symmetric weight matrices.

$$W_1 = \begin{pmatrix} \begin{matrix} 0&0&0&1 \\ 0&0&1&0 \\ 0&1&0&0 \\ 1&0&0&0 \end{matrix} & \bigcirc \\ \bigcirc & \begin{matrix} 0&0&0&1 \\ 0&0&1&0 \\ 0&1&0&0 \\ 1&0&0&0 \end{matrix} \end{pmatrix}$$

$$W_2 = \begin{pmatrix} \begin{matrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{matrix} & \bigcirc \\ \bigcirc & \begin{matrix} 0&1&0&0 \\ 1&0&0&0 \\ 0&0&0&1 \\ 0&0&1&0 \end{matrix} \end{pmatrix}$$

Examples for symmetric weight matrices with row sums equal to one.

$$W_3 = \frac{1}{2} \begin{pmatrix} \begin{matrix} 0&1&0&0&1 \\ 1&0&1&0&0 \\ 0&1&0&1&0 \\ 0&0&1&0&1 \\ 1&0&0&1&0 \end{matrix} & \bigcirc \\ \bigcirc & \begin{matrix} 0&1&0&0&1 \\ 1&0&1&0&0 \\ 0&1&0&1&0 \\ 0&0&1&0&1 \\ 1&0&0&1&0 \end{matrix} \end{pmatrix}.$$

$$W_4 = \frac{1}{4} \begin{pmatrix} \begin{matrix} 0&1&1&1&1 \\ 1&0&1&1&1 \\ 1&1&0&1&1 \\ 1&1&1&0&1 \\ 1&1&1&1&0 \end{matrix} & \bigcirc \\ \bigcirc & \begin{matrix} 0&1&1&1&1 \\ 1&0&1&1&1 \\ 1&1&0&1&1 \\ 1&1&1&0&1 \\ 1&1&1&1&0 \end{matrix} \end{pmatrix}.$$