The Efficiency of OLS Estimator in the Linear Regression Model With Spatially Correlated Errors

Butte Gotu[1]

Department of Statistics, University of Dortmund

Vogelpothsweg 87, 44221 Dortmund, Germany

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

Key words: Ordinary least squares, Generalized least squares, Efficiency, Spatial error process, Spatial correlation.

1 Introduction

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

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

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 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 $Cov(\tilde{\beta}) = \sigma_\epsilon^2(X'V_*^{-1}X)^{-1}$.

[1] This work was partly supported by the Deutsche Forschungsgemeinschaft (DFG), Graduiertenkolleg “Angewandte Statistik”.

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 e.g. 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 and the degree of correlation. 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 are constructed by using the measure 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 $\hat{\beta}$.

2 Bounds for the relative Efficiency of OLS Estimator

In order to analyse the efficiency of OLS relative to GLS estimator, one needs the structure of the covariance matrix of the disturbance vector u. So, we start by specifying stationary 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

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

where ρ denotes a spatial correlation coefficient and ϵ 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 nonnegative weights defined by (see Cliff and Ord, 1981, pp. 17-19)

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

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

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

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

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

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

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

To ensure that V_* is positive definite, the possible values of ρ must be identified (see Horn and Johnson, 1985, p. 301). According to the assumptions given in model (1) the matrix V_* has identical diagonal elements, and denoting this element by υ, we get

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

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

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

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

$$\begin{aligned} e_1(\rho) &:= \frac{1}{2}\,||P_X V \ - \ V P_X||^2 \\ &= \frac{1}{2}\,tr((P_X V \ - \ V P_X)'(P_X V \ - \ V P_X)) \\ &= tr(P_X V^2) \ - \ tr\,(P_X V)^2 \quad . \end{aligned} \tag{8}$$

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 $e_1(\rho) \neq 0$.

Let $\mu_i(A)$ denote the i-th eigenvalue of a $T \times T$ matrix A. Under the assumptions that $X'X = I$, V positive definite and $T \geq 2k$, Bloomfield and Watson give the following upper bound for $e_1(\rho)$:

$$e_1(\rho) \quad \leq \quad \frac{1}{4}\sum_{i=1}^{k}(\mu_i(V) - \mu_{T-i+1}(V))^2 \quad , \tag{9}$$

where the eigenvalues of V are in ascending order.

Remarks:

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 (9) will be large.

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

$$y = \tilde{X}\delta + u \tag{10}$$

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 (10).

By inserting $V = (1/\upsilon)V_*$ in (8), we have

$$e_1(\rho) = \frac{1}{\upsilon^2}\{tr(P_X V_*^2) \; - \; tr\,(P_X V_*)^2\}. \tag{11}$$

Using the result of Bloomfield and Watson (1975), under the assumption that $X'X = I$, V_* positive definite and $T \geq 2k$, and applying (11) we obtain

$$e_1(\rho) \quad \leq \quad \frac{1}{4\upsilon^2}\sum_{i=1}^{k}(\mu_i(V_*) - \mu_{T-i+1}(V_*))^2 \quad , \tag{12}$$

where the eigenvalues of V_* are in ascending order.

In the following the upper bounds of $e_1(\rho)$ will be given, by applying the relationship given in (12) 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

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

Proof:

For a conditional spatial MA(1) process 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,

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

where the eigenvalues $\mu_i(V_*)$ and $\mu_i(W)$, $i = 1, \cdots, T$ are in ascending order. Inserting (14) in (12) completes the proof. $\diamond$

Remarks:

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 difference within the pairs of eigenvalues of W are large.

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

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).

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

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

Proof:

MA(1) process:

Under a spatial MA(1) process we have

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

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

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

implying $\upsilon = 1 + \rho^2$ and $\mu_i(V_*) = (1 + \rho^2) + 2\rho\mu_i(W)$. Inserting these eigenvalues in (12) we get

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

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

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

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

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

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

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

where the eigenvalues are in ascending order. Inserting (15) in (12) and using the fact that $\mu_i(W) \in \{-1, 1\}$ completes the proof. $\diamond$

Remark:

If the diagonal elements of V_* are not identical, then (9) holds when V_* is used instead of V.

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

Theorem 1

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

Proof:

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

$$\begin{aligned} e_1(\rho) &= tr(P_X V^2) - tr\,(P_X V)^2 = tr(P_X V(V - P_X V)) \\ &= tr(P_X V M_X V) \quad . \end{aligned}$$

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

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

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

If the ratio of the mean squared errors are 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 2

Let $X^{'}X = I$. Then

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

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

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

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

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

The second inequality in (19) implies

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

Using (17) to (19) we have

$$\begin{aligned} tr\,(Cov(X\tilde{\beta})) &\geq \sum_{i}^{k} \frac{1}{\mu_{T-k+i}(V^{-1})} \\ &= \sum_{i}^{k} \mu_i(V) \\ tr\,(Cov(X\hat{\beta})) &\leq \sum_{i=1}^{k} \mu_{T-k+i}(V) \quad . \end{aligned} \tag{20}$$

From (20) it is clear that

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

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

Remark:

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 be small, but never less than the ratio of the smallest and the largest eigenvalues $\mu_{min}(V)/\mu_{max}(V)$.

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

Corollary 3

Assume that the matrix X fulfills $X'X = I$. 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

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

Proof:

MA(1) process

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

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

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

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

so that applying Theorem 2 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:

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

and (21) follows by applying Theorem 2. $\diamond$

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.

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

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

where $|\cdot|$ 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 $T \times k$ matrices and assume that $B'B$ is nonsingular. 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, 1988, 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 3

Let $X^{'}X = I$. Then

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

Proof:

By applying Poincaré separation theorem we get

$$\begin{aligned}
\prod_{i=1}^{k} \mu_i(V) &\leq \prod_{i=1}^{k} \mu_i(X^{'}VX) \leq \prod_{i=1}^{k} \mu_{T-k+i}(V) \\
\prod_{i=1}^{k} \mu_i(V^{-1}) &\leq \prod_{i=1}^{k} \mu_i(X^{'}V^{-1}X) \leq \prod_{i=1}^{k} \mu_{T-k+i}(V^{-1}) \quad ,
\end{aligned}$$

where the eigenvalues are in ascending order. This implies

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

so that

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

Furthermore,

$$\begin{aligned}
|X^{'}VX| &\leq \prod_{i=1}^{k} \mu_{T-k+i}(V) \\
|X^{'}V^{-1}X| &\leq \prod_{i=1}^{k} \mu_{T-k+i}(V^{-1}) \quad .
\end{aligned}$$

This implies

$$\begin{aligned} \frac{1}{|X'VX|} &\geq \prod_{i=1}^{k} \frac{1}{\mu_{T-k+i}(V)} \\ |X^{'}V^{-1}X| &\leq \prod_{i=1}^{k} \frac{1}{\mu_i(V)} \quad . \end{aligned} \tag{25}$$

According to the definition, we have

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

and using (25) yields the asserted result. ◇

<u>*Remark:*</u>

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.

<u>**Corollary 4**</u>

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

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

<u>**Proof:**</u>

The proof follows by applying Theorem 3 using the bounds of the eigenvalues of the matrix V given in (22) and (23). ◇

<u>*Remark:*</u>

When the diagonal elements of V_* are not identical, meaning that the $u_i^{'}s$ have different variances, we get (see Theorems 2 and 3)

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

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

Acknowledgements:

The author is grateful to Prof. Dr. S. Schach, Prof. Dr. G. Trenkler and Dr. J. Groß for their useful comments on earlier drafts of the paper.

References

Basilevsky, A. (1983): *Applied matrix algebra in the statistical sciences.* Elsevier Science Publishing, New York.

Bloomfield, P. and Watson, G. S. (1975): " The Inefficiency of Least Squares ", *Biometrika* 62, 121-128.

Cliff, A. D. und Ord, J. K. (1981): *Spatial processes: Models and applications.* Pion, London.

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.

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

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

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

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

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

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.

Krämer, W. (1984): "High Correlation among Errors and the Efficiency of Ordinary Least Squares in the Linear Models", *Statistical Papers* 25, 135-142.

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.

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.

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

Watson, G. S.(1955): " Serial Correlation in Regression Analysis", *Biometrika* 42, 327-341.