The Equality of OLS and GLS Estimators in the Linear Regression Model When the Disturbances are Spatially Correlated

Butte Gotu[1]

Department of Statistics, University of Dortmund
Vogelpothsweg 87, 44221 Dortmund, Germany

Abstract:
Necessary and sufficient conditions for the equality of ordinary least squares and generalized least squares estimators in the linear regression model with first-order spatial error processes are given.

Key words: Ordinary least squares, Generalized least squares, Best linear unbiased estimator, Spatial error process, Spatial correlation.

1 Introduction

Consider the linear regression model for spatial correlation

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

where y is a $T \times 1$ observable random vector, X is a $T \times k$ matrix of known constants with full column rank k, β is a $k \times 1$ vector of unknown parameters, ϵ is a $T \times 1$ random vector with expectation zero and covariance matrix $Cov(\epsilon) = \sigma_\epsilon^2 I$ (I is the T-dimensional identity matrix and σ_ϵ^2 an unknown positive scalar). C denotes a $T \times T$ matrix such that the product CC' is positive definite.

The ordinary least squares (OLS) and the generalized least squares (GLS)

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

estimators of the vector of unknown parameters 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 $Cov(\hat{\beta}) = \sigma_\epsilon^2(X'X)^{-1}X'V_*X(X'X)^{-1}$, $Cov(\tilde{\beta}) = \sigma_\epsilon^2(X'V_*^{-1}X)^{-1}$, where $V_* = CC'$.

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. Since $Cov(u)$ usually involves unknown parameters like spatial correlation coefficient, it is natural to ask when both estimators coincide so that the OLS estimator can be applied without loss of efficiency. Many of the criteria developed for the purpose of checking the equality of least squares estimators are not operational because of the unknown parameters involved (see Puntanen and Styan, 1989).

In this paper, conditions under first-order spatial error processes which can be verified in practice by using spatial weights matrix with known nonnegative weights and the matrix X of known constants are developed. The first group of conditions is based on the invariance property of the column space of the matrix X under V_* (Kruskal, 1968), whereas the second one uses the symmetry of the product $P_X V_*$ (Zyskind, 1967), with $P_X = X(X'X)^{-1}X'$.

2 Equality of OLS and GLS estimators

In assessing the conditions for the equality of OLS and GLS estimators, the structure of the covariance of the disturbance vector u plays an important role. So, we start by giving possible structures of $Cov(u)$ under first-order spatial error processes.

Let the components of u follow a first-order spatial autoregressive (AR(1)) process

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

or, in matrix form

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

where ρ denotes a spatial correlation coefficient for a given area partitioned into T nonoverlapping regions R_i, $i = 1, \cdots, T$. W is a weights matrix with known nonnegative weights defined by (see Cliff and Ord, 1981, pp. 17-19)

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

The element w_{ij} of the weights matrix indicates the strength of the effect of region R_j on region R_i. Under first-order spatial moving average (MA(1)) process the components of u follow the pattern

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

or, in matrix form

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

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

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

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

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

Note that the possible values of ρ must be identified to ensure that V_* is positive definite.

In the following we investigate conditions for the equality of OLS and GLS estimators by applying the result: two unbiased estimators coincide almost surely if and only if their covariances are equal (see Puntanen and Styan, 1989, p. 154). This means, OLS and GLS are equal if and only if their covariances are equal.

Let $\mathcal{R}(X)$ denote a k-dimensional space spanned by the columns of X. The well known Kruskal's (1968) column space condition for the equality of OLS estimator $\hat{\beta}$ and GLS estimator $\tilde{\beta}$ in model (1) states that both estimators coincide if and only if

$$\mathcal{R}(V_* X) = \mathcal{R}(X) \quad , \tag{6}$$

where V_* is assumed to be a nonsingular matrix.

In order to apply Kruskal's condition, the value of the unknown parameter ρ in the Matrix V_* must be given in addition to X. In practice ρ typically will be unknown and one needs a more applicable condition to check the equality. Based on Kruskal's theorem Krämer and Donninger (1987) give a sufficient condition which can be verified in practice when the disturbances follow a first-order spatial autoregressive process. Baksalary (1988) generalizes this result for first-order spatial error processes as follows.

Theorem 1

Let W be a $T \times T$ weights matrix and V_* be a $T \times T$ positive definite matrix of the form

$$V_* = (I + \rho W')(I + \rho W) \quad \text{or} \quad V_* = (I + \rho W)(I + \rho W') \quad ,$$

where $\rho \neq 0$ is a scalar. If $\mathcal{R}(WX) \subseteq \mathcal{R}(X)$ and $\mathcal{R}(W'X) \subseteq \mathcal{R}(X)$, then $\hat{\beta} = \tilde{\beta}$.

Proof:

The conditions $\mathcal{R}(WX) \subseteq \mathcal{R}(X)$ and $\mathcal{R}(W'X) \subseteq \mathcal{R}(X)$ imply that

$$\mathcal{R}((I + \rho W)X) = \mathcal{R}(X) \quad \text{and} \quad \mathcal{R}((I + \rho W')X) = \mathcal{R}(X)$$

irrespective of ρ. From this we get

$$\mathcal{R}(V_* X) = \mathcal{R}((I + \rho W')(I + \rho W)X) = \mathcal{R}(X)$$

and the equality of the estimators follows from Kruskal's theorem. ◇

The following sufficient condition for the equality under a specific matrix V_* is also based on condition (6).

Theorem 2

Let $\mathbf{b}_1$ and $\mathbf{b}_2$ be $T \times 1$ vectors, and let V_* be a $T \times T$ positive definite matrix of the pattern

$$V_* = cI + \mathbf{b}_1\,\mathbf{b}_2^{'} + \mathbf{b}_2\,\mathbf{b}_1^{'}$$

with a scalar c. If $\mathbf{b}_1 \in \mathcal{R}(X)$ and $\mathbf{b}_2 \in \mathcal{R}(X)$, then $\mathcal{R}(V_* X) = \mathcal{R}(X)$.

Proof: See Mathew, 1984, pp. 207-208. ◇

By combining the results in theorems 1 and 2 the following sufficient condition for the equality of OLS and GLS estimators can be formulated.

Corollary 1

Let $\mathbf{d}$ be a $T \times 1$ vector and V_* be a $T \times T$ positive definite matrix of the pattern

$$V_* = c_1 I + c_2\,W^* + c_3\,\mathbf{d}\,\mathbf{d}^{'} \quad ,$$

where c_1, c_2, c_3 are scalars, and W^* is a $T \times T$ matrix. If $\mathcal{R}(W^* X) \subseteq \mathcal{R}(X)$ and $\mathbf{d} \in \mathcal{R}(X)$, then $\hat{\beta} = \tilde{\beta}$.

Proof: The proof follows from Theorems 1 and 2. ◇

Simple examples show that the conditions of the above results are not necessary for the equality of OLS and GLS estimators (see Baksalary, 1988 and Gotu, 1997). The theorem below, based on the result given by Baksalary (1988), provides necessary and sufficient conditions.

Theorem 3

Let W be a $T \times T$ weights matrix and V_* be a $T \times T$ matrix of the form

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

Further, let ϕ be given by: $\phi = \{\rho \neq 0 : V_*$ positive definite and $|\rho| < 1\}$. Then the following conditions are equivalent:

(i) $\mathcal{R}(V_* X) = \mathcal{R}(X)$ for all $\rho \in \phi$.

(ii) $\mathcal{R}(V_* X) = \mathcal{R}(X)$ for two different $\rho_1, \rho_2 \in \phi$.

(iii) $\mathcal{R}((W + W')X) \subseteq \mathcal{R}(X)$ and $\mathcal{R}(W'WX) \subseteq \mathcal{R}(X)$.

Proof:

(i) $\Longrightarrow$ (ii):

The condition $\mathcal{R}(V_* X) = \mathcal{R}(X)$ for $\rho \neq 0$ holds if and only if

$$\mathcal{R}((W + W' + \rho W W')X) \subseteq \mathcal{R}(X). \tag{7}$$

If (7) is valid for all $\rho \in \phi$, then

$$\begin{aligned} \mathcal{R}((W + W' + \rho_1 W W')X) &\subseteq \mathcal{R}(X) \\ \mathcal{R}((W + W' + \rho_2 W W')X) &\subseteq \mathcal{R}(X). \end{aligned} \tag{8}$$

(ii) $\Longrightarrow$ (iii):

From equation (8) we get $\mathcal{R}((\rho_1 - \rho_2)WW'X) \subseteq \mathcal{R}(X)$. This implies $\mathcal{R}(WW'X) \subseteq \mathcal{R}(X)$, and $\mathcal{R}((W' + W)X) \subseteq \mathcal{R}(X)$ follows from (7).

(iii) $\Longrightarrow$ (i): Follows direct from (7). $\diamond$

Remarks:

- *The matrix V_* is positive definite if $I + \rho W$ is nonsingular and the nonsingularity of $I + \rho W$ holds if there exists a matrix-norm which satisfies the inequality $|\rho|\,||W|| < 1$ (see Horn and Johnson, 1985, p. 301). For any given weights matrix W with row sums equal to one, the maximum row sum matrix-norm is equal to one, so the matrix $I + \rho W$ is nonsingular for $|\rho| < 1$.*

- *Let A be a symmetric matrix. Then $\mathcal{R}(AX) \subseteq \mathcal{R}(X)$ if and only if $P_X A = AP_X$. This means condition (iii) is equivalent to $(W+W')P_X = P_X(W+W')$ and $WW'P_X = P_X WW'$.*

- *Theorem 3 applies also for V_* matrix of the form*

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

  because $\qquad \mathcal{R}(V_*X) = \mathcal{R}(X) \quad \Longleftrightarrow \quad \mathcal{R}(V_*^{-1}X) = \mathcal{R}(X)\,.$

- *If OLS and GLS estimators are equal for two different values of ρ, that is $\mathcal{R}(V_*X) = \mathcal{R}(X)$ for different $\rho_1, \rho_2 \in \phi$, then from the equivalence of (i) and (ii) follows that both estimators are equal for all $\rho \in \phi$.*

- *Condition (iii) can be applied to check the equality of OLS and GLS without specifying the value of ρ.*

- *For V_* matrix of the form $(I-\rho W)^{-1}$ or $I+\rho W$, where W is symmetric, condition (iii) should be restated as $\mathcal{R}(WX) \subseteq \mathcal{R}(X)$.*

- *Let W_1 and W_2 be $T \times T$ weights matrices, and D_1 and D_2 be $T \times T$ diagonal matrices with full rank. Suppose that $W_1'D_1^{-1} = D_1^{-1}W_1$ and $D_2W_2' = W_2D_2$. If V_* is of the pattern $(I-\rho W_1)^{-1}D_1$ or $(I+\rho W_2)D_2$, condition (iii) should, accordingly, be restated as*
  $\mathcal{R}(D_1^{-1}X) \subseteq \mathcal{R}(X)$ and $\mathcal{R}(D_1^{-1}W_1X) \subseteq \mathcal{R}(X)$;
  $\mathcal{R}(D_2X) \subseteq \mathcal{R}(X)$ and $\mathcal{R}(D_2W_2X) \subseteq \mathcal{R}(X)$.

In the following, conditions for the equality of least squares estimators for a subvector of β will be discussed.
Suppose that X_1 and X_2 are submatrices of X, and β_1 and β_2 be subvectors of β. Further, let $\hat{\beta}_2$ and $\tilde{\beta}_2$ be the respective subvectors of $\hat{\beta}$ and $\tilde{\beta}$. Splitting model (1) into

$$y = X_1\beta_1 + X_2\beta_2 + u \quad ,$$

Krämer et al. (1996) give the following necessary and sufficient condition for the equality of $\hat{\beta}_2$ and $\tilde{\beta}_2$:

$$\hat{\beta}_2 = \tilde{\beta}_2 \quad \Longleftrightarrow \quad \mathcal{R}(V_* X^\perp) \subseteq (\mathcal{R}(X_1) \oplus \mathcal{R}(X^\perp)) \quad ,$$

where $X^\perp$ is a matrix such that $\mathcal{R}(X^\perp) = \mathcal{R}(X)^\perp$, the orthogonal complement of $\mathcal{R}(X)$, and $\oplus$ is the direct sum of subspaces.
The problem with the above condition is, as in Kruskal's theorem, that the unknown parameter ρ in the matrix V_* should be given. The following result, which is based on Theorem 3, provides a necessary and sufficient condition for the equality of $\hat{\beta}_2$ and $\tilde{\beta}_2$ under the first-order spatial error process that works without specifying the value of ρ.

Corollary 2

Let W be $T \times T$ weights matrix and V_* be a $T \times T$ matrix of the form

$$V_* = (I + \rho W)(I + \rho W') \quad , \tag{9}$$

where $\rho \in \phi$. Then the following statements are equivalent:

(a) $\mathcal{R}(V_* X^\perp) \subseteq \mathcal{R}(X_1) \oplus \mathcal{R}(X^\perp)$ for all $\rho \in \phi$.

(b) $\mathcal{R}(V_* X^\perp) \subseteq \mathcal{R}(X_1) \oplus \mathcal{R}(X^\perp)$ for two different $\rho_1, \rho_2 \in \phi$.

(c) $\mathcal{R}((W + W')X^\perp) \subseteq \mathcal{R}(X_1) \oplus \mathcal{R}(X^\perp)$ and $\mathcal{R}(WW'X^\perp) \subseteq \mathcal{R}(X_1) \oplus \mathcal{R}(X^\perp)$.

Proof: See Theorem 3. $\diamond$

Remarks:

- *In order to check the equality of $\hat{\beta}_2$ and $\tilde{\beta}_2$, statement (iii) can be applied independent of ρ.*

- *For the matrix of the form (9) the following holds (see Theorem 3): If*

  $$\mathcal{R}(WX^{\perp}) \subseteq \mathcal{R}(X_1) \oplus \mathcal{R}(X^{\perp}) \quad \textit{and} \quad \mathcal{R}(W'X^{\perp}) \subseteq \mathcal{R}(X_1) \oplus \mathcal{R}(X^{\perp}),$$

  then $\hat{\beta}_2 = \tilde{\beta}_2$.

Another well known condition for the coincidence of OLS and GLS estimators in the linear regression model (1) is based on the symmetry of the matrix product $P_X V_*$. That is, in the regression model (1)

$$\hat{\beta} = \tilde{\beta} \quad \Longleftrightarrow \quad P_X V_* = V_* P_X \quad . \tag{10}$$

For the application of this condition the values of the unknown parameters in the matrix V_* should again be given. The following sufficient condition can be applied under the first-order spatial error processes, irrespective of the parameters in V_*.

Corollary 3

Assume that the components of the disturbance vector u in model (1) follow a first-order spatial moving average or autoregressive process. Let W be a $T \times T$ weights matrix. The estimators $\hat{\beta}$ and $\tilde{\beta}$ coincide if

$$P_X W = W P_X \quad . \tag{11}$$

Proof:

MA(1) process:

Under spatial MA(1) error process the matrix V_* is given by

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

From equation (10) the estimators $\hat{\beta}$ and $\tilde{\beta}$ coincide if and only if

$$P_X V_* = V_* P_X \quad .$$

The above equation holds if for $\rho \neq 0$

$$P_X W' + P_X W + \rho P_X W W' = W' P_X + W P_X + \rho W W' P_X \quad . \tag{12}$$

By equation (11), applying the symmetry of P_X, we get $P_X W' = W' P_X$ and from (12) follows $P_X V_* = V_* P_X$ implying the equality of the estimators $\hat{\beta}$ and $\tilde{\beta}$.

AR(1) process:

Under spatial AR(1) error process we have

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

Furthermore, $P_X V_* = V_* P_X$ if and only if

$$P_X V_*^{-1} = V_*^{-1} P_X \quad . \tag{13}$$

Equation (13) holds if

$$\rho P_X W' W - P_X W' - P_X W = \rho W' W P_X - W' P_X - W P_X \tag{14}$$

with $\rho \neq 0$. By equation (14), applying the symmetry of P_X and equation (11), we obtain $P_X V_*^{-1} = V_*^{-1} P_X$ implying the equality of $\hat{\beta}$ and $\tilde{\beta}$. $\diamond$

Remarks:

It can be shown that the condition of Corollary 3 is also necessary if (see Gotu, 1997)

- *the weights matrix W is symmetric and orthogonal.*
- *the components of the disturbance vector u follow a conditional first-order spatial process with V_* given in (5).*

A counter-example that the condition of Corollary 3 is necessary in general can be obtained by taking

$$W = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{pmatrix} \qquad X = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$V_* = (I + \rho W')(I + \rho W)$ and $\rho = 3/4$. In this case $P_X V_* = V_* P_X$ although $P_X W \neq W P_X$.

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

Baksalary, J. K. (1988): "Two Notes of JKB", *unpublished.*

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

Gotu, B. (1997): "Schätz- und Testmethoden im linearen Regressionsmodell bei Vorliegen von räumlicher Korrelation", *Doctoral Thesis, University of Dortmund.*

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

Krämer, W., Bartels, R. and Fiebig, D. G. (1996): " Another Twist on the Equality of OLS and GLS", *Statistical Papers* 37, 277-281.

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.

Kruskal, W. (1968): "When are Gauss-Markov and Least Squares Estimator Identical? A Coordinate-Free Approach", *The Annals of Mathematical Statistics* 39, 70-75.

Mathew, T. (1984): "On Inference in a General Linear Model with an Incorrect Disperssion Matrix ", in Calinski, T. and Klonecki, W. (eds.): *Linear Statistical Inference, Proceedings of Poznan Conference.* Springer Verlag, Berlin, 200-210.

Puntanen, S. and Styan, G. P. H. (1989): "The Equality of the Ordinary Least Squares Estimator and the Best Linear Unbiased Estimator", *The American Statistician* 43, 153-164.

Zyskind, G. (1967): " On Canonical Forms, Non-Negative Covariance Matrices and Best and Simple Least Squares Linear Estimators in Linear Models", *The Annals of Mathematical Statistics* 38, 1092-1109.