$+ y^* = 0$, violating the requirement that $\Pr(X + y_1$ and $X + y_2$ have opposite signs$) = 0$. For symmetric distributions, it is necessary (and easily sufficient) for $E|X + Y| = E|X|$ to hold that $\Pr(|X| \geq |Y|) = 1$.

The symmetry assumption is relevant. Consider $X = 2$ or -4 with probabilities $\frac{2}{3}$ and $\frac{1}{3}$, $Y = 3$ or -1 with probabilities $\frac{1}{4}$ and $\frac{3}{4}$. $E|X + Y| = E|X| = \frac{8}{3}$. For nonsymmetric distributions with $y_{\min}$ = smallest (most negative) possible Y value, $y_{\max}$ = largest possible Y value, and say, $y_{\max} > -y_{\min}$, the necessary and sufficient condition for equality is $\Pr(y_{\min} < X < -y_{\min}) = 0$; a similar condition holds if $y_{\max} < -y_{\min}$.

In Lord's example, X and Y were random errors around an unknown parameter value. To obtain Lord's paradox, with symmetric error distributions, it is necessary that the X error *always* be larger in magnitude (or at least no smaller) than the Y error. In particular, the paradox cannot occur if there is any probability at all of a small (relative to Y) X error. One hopes that cases in which a random error term is guaranteed to be large in magnitude are rare in practice.

If we make the additional assumption that X and Y are identically distributed, we have that $E|X + Y| = E|X| = E|Y|$ if and only if $\Pr(|X| \geq |Y|) = 1$ and $\Pr(|Y| \geq |X|) = 1$, that is, if and only if $\Pr(|X| = |Y|) = 1$. The only way to have independent, nondegenerate X and Y with $\Pr(|X| = |Y|) = 1$ and $E(X)$ and $E(Y) = 0$ is to have $\Pr(X = c) = \Pr(X = -c) = \Pr(Y = c) = \Pr(Y = -c) = .5$. Lord's example was precisely this case, with $c = 1$.

[Received December 1983. Revised March 1984.]

REFERENCES

LORD, F.M. (1983), "MAD Query," *The American Statistician*, 37, 343.

MOOD, A.M., GRAYBILL, F.A., and BOES, D.C. (1974), *Introduction to the Theory of Statistics* (3rd ed.), New York: McGraw-Hill.

Nonnegative Definiteness of the Sample Autocovariance Function

A. IAN McLEOD and CARLOS JIMÉNEZ*

Various textbooks on time series analysis assert that the usual version of the sample autocovariance function (1) is nonnegative definite. Two simple proofs of this result are presented.

KEY WORDS: Autocovariance function; Time domain.

1. INTRODUCTION

The theoretical autocovariance function γ_k $(k = 0, \pm 1, \pm 2, \ldots)$ of a stationary time series may be estimated from n consecutive observations, z_t $(t = 1, \ldots, n)$, by the sample autocovariance function

$$c_k = \sum_{t=k+1}^{n} (z_t - \bar{z})(z_{t-k} - \bar{z})/n, \qquad k \geq 0, \qquad (1)$$

where $\bar{z} = \Sigma z_t/n$. For $k < 0$, $c_k = c_{-k}$. Note that $c_k = 0$ for $k \geq n$, since by convention the empty sum is zero. One of the reasons this estimate is preferred (Jenkins and Watts 1969, p. 184; Fuller 1976, p. 236) to alternatives such as dividing by $(n - k)$ rather than n is that the resulting estimate of the covariance matrix, C_m, of m $(m \geq 1)$ consecutive observations,

$$C_m = (c_{i-j})_{m \times m}, \qquad (2)$$

is a nonnegative definite matrix. A sequence c_k for which the resulting matrices C_m in (2) are nonnegative definite for $m > 0$ is said to be nonnegative definite. It is easily seen that γ_k $(k = 0, \pm 1, \pm 2, \ldots)$ is nonnegative definite (Box and Jenkins 1976, p. 28). To our knowledge, however, the only proof for c_k available in the literature (e.g., see Priestley 1981, p. 323) uses properties of the periodogram. This method is fairly indirect and not elementary, and it is particularly undesirable in an elementary course oriented toward the time domain. Two simple proofs of the nonnegative definiteness of c_k are given in the next sections.

2. ALGEBRAIC PROOF

Without loss of generality, let $\bar{z} = 0$. Let $z_t = 0$ for $t < 1$ or $t > n$ and define the column vector $Z_i = (z_{i+1}, \ldots, z_{i+n})'$, $-n < i < n$. Then

$$nC_n = \sum_{i=1-n}^{n-1} Z_i Z_i' \qquad (3)$$

and $\alpha' Z_i Z_i' \alpha = (Z_i' \alpha)'(Z_i' \alpha) \geq 0$ for any column vector α.

Thus C_n is the sum of nonnegative definite matrices and consequently must also be nonnegative definite. It follows that C_k is nonnegative definite for any $k > 0$.

3. MOVING AVERAGE PROOF

Let $h_t (t = 1, 2, \ldots)$ be a moving average process of order n defined by

*A. Ian McLeod is Associate Professor and Carlos Jiménez is a Ph.D. candidate, both at the Department of Statistical and Actuarial Sciences, The University of Western Ontario, London, Ontario, Canada N6A 5B9.

$$h_t = \sum_{i=1}^{n} z_i e_{t-i}, \qquad (4)$$

where e_t is a white noise sequence with variance $1/n$. It is easily verified that the theoretical autocovariance function of h_t is just the sample autocovariance function of $z_1, \ldots, z_n$ given by (1). Consequently, c_k is also nonnegative definite.

[*Received July 1983. Revised February 1984.*]

REFERENCES

BOX, G.E.P., and JENKINS, G.M. (1976), *Time Series Analysis Forecasting and Control* (2nd ed.), San Francisco: Holden-Day.

FULLER, W.A. (1976), *Introduction to Statistical Time Series*, New York: John Wiley.

JENKINS, G.M., and WATTS, D.G. (1969), *Spectral Analysis and Its Applications*, San Francisco: Holden-Day.

PRIESTLEY, M.B. (1981), *Spectral Analysis and Time Series, Vol. 1: Univariate Series*, London: Academic Press.

On Necessary and Sufficient Conditions for Ordinary Least Squares Estimators to Be Best Linear Unbiased Estimators

GEORGE A. MILLIKEN and MOHAMMED ALBOHALI*

Two often-quoted necessary and sufficient conditions for ordinary least squares estimators to be best linear unbiased estimators are described. Another necessary and sufficient condition is described, providing an additional tool for checking to see whether the covariance matrix of a given linear model is such that the ordinary least squares estimator is also the best linear unbiased estimator. The new condition is used to show that one of the two published conditions is only a sufficient condition.

KEY WORDS: Linear model; Ordinary least squares; Covariance matrix.

1. INTRODUCTION

The general linear model with a full rank design matrix is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \qquad (1.1)$$

where $\mathbf{X}$ is an $n \times p$ matrix of rank p, $E(\boldsymbol{\epsilon}) = \mathbf{0}$, var $(\boldsymbol{\epsilon}) = \boldsymbol{\Sigma}$, and $\boldsymbol{\Sigma}$ is positive definite. The ordinary least squares estimator (OLSE) of $\boldsymbol{\beta}$ is $\hat{\boldsymbol{\beta}}_{\text{LS}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$, which is chosen to minimize the quantity $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$. The best linear unbiased estimator (BLUE) of $\boldsymbol{\beta}$ is chosen to minimize the quantity $(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$, and when $\boldsymbol{\Sigma}$ is known, the BLUE of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}}_{\text{BL}} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}.$$

Several authors have discussed conditions for $\boldsymbol{\Sigma}$ when it is unknown, for which $\hat{\boldsymbol{\beta}}_{\text{LS}}$ is also the BLUE. Graybill (1976, Theorem 6.8.1) stated that $\hat{\boldsymbol{\beta}}_{\text{LS}}$ is also BLUE if and only if there exists a nonsingular matrix $\mathbf{F}$ such that $\boldsymbol{\Sigma}\mathbf{X} = \mathbf{X}\mathbf{F}$. McElroy (1967) considered the case in which the linear model has an intercept, that is, one of the columns of $\mathbf{X}$ is a vector of 1's. For the intercept linear model, McElroy's states that $\hat{\boldsymbol{\beta}}_{\text{LS}}$ is also BLUE if and only if $\boldsymbol{\Sigma} = \sigma^2[q\mathbf{I} + (1 - q)\mathbf{J}]$, where $\mathbf{I}$ is an $n \times n$ identity matrix, $\mathbf{J}$ is an $n \times n$ matrix of 1's, and $|q| < 1$. McElroy's condition is often quoted and was used most recently by Lowerre (1983). Unfortunately McElroy's condition is only sufficient. The theorem in the next section presents a necessary and sufficient condition (NSC), which can easily be used to show that McElroy's condition is not necessary.

*George A. Milliken is Professor, Department of Statistics, Kansas State University, Manhattan, KS 66506. Mohammed Albohali is Assistant Professor, Department of Statistics, Temple University, Philadelphia, PA 19401.

2. ANOTHER NSC

Graybill's condition depends on the existence of the nonsingular matrix $\mathbf{F}$, which may be difficult to find for some choices of $\mathbf{X}$ and $\boldsymbol{\Sigma}$. If an expression for $\boldsymbol{\Sigma}^{-1}$ can be obtained, the following theorem presents another NSC.

Theorem. For the linear model in (1.1), $\hat{\boldsymbol{\beta}}_{\text{LS}}$ is BLUE if and only if

$$\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\mathbf{I} - \mathbf{W}) = \mathbf{0},$$

where $\mathbf{W} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

Proof. In order for $\hat{\boldsymbol{\beta}}_{\text{LS}}$ to be BLUE, $\mathbf{X}$ and $\boldsymbol{\Sigma}$ must be such that $\hat{\boldsymbol{\beta}}_{\text{LS}} = \hat{\boldsymbol{\beta}}_{\text{BL}}$ or

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y}.$$

Let $\mathbf{U}$ denote the difference

$$\mathbf{U} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{y},$$

and express $\mathbf{y} = \mathbf{W}\mathbf{y} + (\mathbf{I} - \mathbf{W})\mathbf{y}$. Then $\mathbf{U}$ reduces to

$$\mathbf{U} = -(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\mathbf{I} - \mathbf{W})\mathbf{y}.$$

Thus $\hat{\boldsymbol{\beta}}_{\text{LS}}$ is also BLUE if and only if $\mathbf{U} = \mathbf{0}$ for all $\mathbf{y}$ or if and only if

$$(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\mathbf{I} - \mathbf{W}) = \mathbf{0}.$$

Since $(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$ is nonsingular, the condition becomes $\hat{\boldsymbol{\beta}}_{\text{LS}}$ is BLUE if and only if

$$\mathbf{X}'\boldsymbol{\Sigma}^{-1}(\mathbf{I} - \mathbf{W}) = \mathbf{0}.$$