Journal of the Korean
Statistical Society
Vol. 25, No. 2, 1996

The Asymptotic Unbiasedness of S^2 in the Linear Regression Model with Dependent Errors [†]

Sangyeol Lee and Young-Won Kim [1]

Abstract

The ordinary least squares estimator of the disturbance variance in the linear regression model with stationary errors is shown to be asymptotically unbiased when the error process has a spectral density bounded from the above and away from zero. Such error processes cover a broad class of stationary processes, including ARMA processes.

Key Words : OLS estimator; Regression model with dependent errors; Unbiased of S^2; Spectral density.

[†]This work was supported by Sookmyung Women's University Research Fund in 1996.
[1] Department of Statistics, Sookmyung Women's University, Yongsan-ku, Seoul, 140-742, Korea

1. INTRODUCTION

Let us consider the linear regression model

$$y_j = \beta_j x_j + \varepsilon_j,\ j = 1, \ldots, n, \tag{1.1}$$

where x_j are $p \times 1$ nonstochastic design vectors and the error $\{\varepsilon_j\}$ is a strictly stationary process with mean zero and variance σ^2. In literature, many researchers have studied the asymptotic properties of the least squares estimator of β in the regression model with dependent errors. See, for example, Solo(1981) and Lai and Wei(1982). In such articles, the error process is assumed to be a stationary process or a sequence of martingale differences.

Although most literatures put their primary interest to the asymptotic behavior of the least squares estimator itself, Sathe and Vinod(1974) and Dunfour(1986) focused on that of S^2, the sum of least squares residuals divided by $n-p$. They showed that if $E\varepsilon\varepsilon' = \sigma^2 V$, where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)'$ and V is a $n \times n$ positive definite matrix, $E\frac{S^2}{\sigma^2}$ is bounded from the below and the above by the mean value of the $n-p$ smallest and largest eigenvlaues of V, respectively. (See also Neudecker(1977, 1978) for related results). Based on the result, Song(1994) established the asymptotic unbiasedness of S^2. The assumed model of Song(1994) was the linear regression model of which the error process is either the first order moving average process or the s-th order autoregressive process.

The objective of this article is to show that under fairly mild condition, the asymptotic unbiasedness for S^2 holds for a broad class of stationary error processes, including ARMA processes. Since the proofs of Song(1994) are hardly applicable to our setting unless the structure of the error process is completely known or the covariance matrix V only depends on a finite number of parameters that can be estimated from data, we here pursue a different approach to obtain the asymptotic unbiasedness result. The main result of this article is addressed in Section 2.

2. MAIN RESULTS

Suppose that the data $(x_1, y_1), \ldots, (x_n, y_n)$ are from the regression model (1.1), and that the error process $\{\varepsilon_j; j \geq 1\}$ is the strictly stationary process with the spectral density f such that

$$0 < m = \inf_{\nu} f(\nu) \leq \sup_{\nu} f(\nu) = M < \infty.$$

Note that a large class of stationary processes satisfy the above assumption. For example, if $\varepsilon_j = \sum_{k=0}^{\infty} a_k \delta_{j-k}$, where $\{\delta_i\}$ is a sequence of iid random variables with mean zero and finite variance, and $\{a_i\}$ is an absolutely summable real sequence with $\sum_{i=0}^{\infty} a_i z^i \neq 0$ for $|z| \leq 1$ in the complex plane, the spectral density of $\{\varepsilon_j\}$ is bounded from the above and away from zero. It is well-known that such processes include the stationary ARMA processes.

Let $\hat{\beta}$ be the least squares estimator of β and let $e_j = y_j - \hat{\beta}' x_j$, $j = 1, \ldots, n$, and $S^2 = \frac{1}{n-p} \sum_{j=1}^{n} e_j^2$. Notice that

$$(n-p)S^2 = \sum_{i=1}^{n} \varepsilon_j^2 - Q_n, \tag{2.1}$$

where

$$Q_n = (\hat{\beta} - \beta)' \left(\sum_{j=1}^{n} x_j x_j' \right) (\hat{\beta} - \beta),$$

which can be rewritten as

$$Q_n = (\sum_{j=1}^{n} x_j \varepsilon_j)' \left(\sum_{j=1}^{n} x_j x_j' \right)^{-1} (\sum_{j=1}^{n} x_j \varepsilon_j). \tag{2.2}$$

The above representation turns out to be very useful to establish the following theorem.

Theorem 1. If $n^{-1} \sum_{j=1}^{n} x_j x_j'$ is nonsingular for all sufficiently large n and converges to a positive definite matrix Σ as n goes to infinity, ES^2 converges to σ^2 as n goes to infinity.

Before we proceed, we introduce a lemma useful to prove Theorem 1.

Lemma 1. Suppose that $\{Y_t; t \geq 1\}$ is a stationary time series of which spectral density g satisfies

$$0 < m = \inf_{\nu} g(\nu) \leq \sup_{\nu} g(\nu) = M < \infty.$$

Then if Γ_n is the covarince matrix of $(Y_1, \ldots, Y_n)$, it follows that

$$2\pi m||z||^2 \leq z'\Gamma_n z \leq 2\pi M||z||^2 \quad \text{for all} \quad z \in R^n.$$

Proof. The lemma follows immediately from the following:

$$z'\Gamma_n z = \int_{-\pi}^{\pi} |\sum_{j=1}^{n} z_j e^{-ij\nu}|^2 f(\nu) d\nu,$$

where $z = (z_1, \ldots, z_n)'$ (cf. Brockwell and Davis (1990, p. 138)).

Proof of Theorem 1. By (2.1) we have

$$\sigma^2 - E\frac{Q_n}{n} \leq \frac{n-p}{n} ES^2 \leq \sigma^2, \tag{2.3}$$

and therefore it suffices to show that $E\frac{Q_n}{n} \to 0$ as $n \to \infty$.

Let

$$A_n = \left\{ (x_1', \ldots, x_n') \in R^{np}; \lambda_{\min}\left(\sum_{j=1}^{n} x_j x_j'\right) < n(\lambda_{min}(\Sigma) - \delta) \right\}, \tag{2.4}$$

where $\lambda_{\min}(E)$ denotes the smallest eigenvalue of any $p \times p$ matrix E and δ is a positive number less than $\lambda_{\min}(\Sigma)$. Then in view of (2.2),

$$\begin{aligned} Q_n &= Q_n I(A_n) + Q_n I(A_n^c) \\ &\leq \sum_{j=1}^{n} \varepsilon_j^2 I(A_n) + \lambda_{\min}^{-1}\left(\sum_{j=1}^{n} x_j x_j'\right) ||\sum_{j=1}^{n} x_j \varepsilon_j||^2 I(A_n^c) \\ &:= I_n + II_n, \end{aligned} \tag{2.5}$$

where $I(\cdot)$ denotes the indicator function. Let $x_j = (x_{1j}, \ldots, x_{pj})'$. Note that $I_n = 0$ for all sufficiently large n, and that II_n is no more than

$$(\lambda_{min}(\Sigma) - \delta)^{-1} \sum_{i=1}^{p} \left(n^{-1/2} \sum_{j=1}^{n} x_{ij}\varepsilon_j\right)^2.$$

By Lemma 1, we have

$$2\pi m\left(n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{n} x_{ij}^2\right) \leq E\sum_{i=1}^{p}\left(n^{-1/2}\sum_{j=1}^{n} x_{ij}\varepsilon_j\right)^2 \leq 2\pi M\left(n^{-1}\sum_{i=1}^{p}\sum_{j=1}^{n} x_{ij}^2\right) \tag{2.6}$$

where $m = \inf_\nu f(\nu)$ and $M = \sup_\nu f(\nu)$. Since for each i, $n^{-1}\sum_{j=1}^n x_{ij}^2$ converges to some constant $c_i > 0$ our assumption, EII_n is bounded uniformly in n and thus $\lim_{n\to\infty} EQ_n/n = 0$. Hence, we establishe the theorem.

3. REMARKS

Note that $\sum_{j=1}^n x_j x_j^{'}$ are nonsingular for all $n \geq n_0$ once $\sum_{j=1}^{n_0} x_j x_j^{'}$ is nonsingular. The condition of $n^{-1}\sum_{j=1}^n x_j x_j^{'}$ converging to a positive definite matrix is a fairly mild condition and can be found in most regression analysis texts.

In view of (2.3)-(2.6), we have that

$$E\frac{S^2}{\sigma^2} \geq \frac{n}{n-p}\Bigg\{1 - (n\sigma^2)^{-1}\sum_{j=1}^n \varepsilon_j^2 I(A_n) + (n\sigma^2)^{-1}(\lambda_{\min}(\Sigma) - \delta)^{-1} \times 2\pi M\left(n^{-1}\sum_{i=1}^p\sum_{j=1}^n x_{ij}^2\right)\Bigg\}.$$

It is possible in real pracice to compute $\lim_{n\to\infty} n^{-1}\sum_{j=1}^n x_j x_j^{'}$ and therefore $\lambda_{\min}(\Sigma)$ for suitably chosen x_j. Let $\delta_m = \lambda_{\min}(\Sigma) - m^{-1}\lambda_{\min}(\Sigma)$ and $m_n = \inf_{m\geq 1}\{m; n \in A_n^c(\delta_m)\}$, where $A_n(\delta_m)$ denotes the set A_n with δ replaced by δ_m. Then $E\frac{S^2}{\sigma^2}$ has a lower bound

$$\frac{n}{n-p}\left\{1 - \frac{2\pi m_n M}{n\sigma^2\lambda_{\min}(\Sigma)}\left(n^{-1}\sum_{i=1}^p\sum_{j=1}^n x_{ij}^2\right)\right\},$$

which suggests as an estimate of the lower bound

$$\frac{n}{n-p}\left\{1 - \frac{2\pi m_n \hat{M}}{nS^2\lambda_{\min}(\Sigma)}\left(n^{-1}\sum_{i=1}^p\sum_{j=1}^n x_{ij}^2\right)\right\},$$

where $\hat{M}$ is a suitable estimate of M that can be obtained from the spectral density estimate based on residuals. (See Chapter 10 of Brocwell and Davis(1990) for the inference of spectral density of time series). Hence, we are able to estimate the lower bound of $E\frac{S^2}{\sigma^2}$. It is, however, impossible to estimate the lower bound through the result of Dunfour(1986) since we do not know the covariance matrix V of $(\varepsilon_1, \ldots, \varepsilon_n)'$. In fact, V cannot be estimated from data unless it only depends on a finite number of parameters. Even worse, the eigenvalues of large matrices are difficult to compute.

One can easily see that Theorem 1 remains valid when the error process is a sequence of martingale differences $\{\varepsilon_j, \mathcal{F}_j; j \geq 0\}$, where $\{\mathcal{F}_j\}$ is a sequence of increasing σ-fields, $E(\varepsilon_j|\mathcal{F}_{j-1}) = 0$ for all j and $\sup_j E(\varepsilon_j^2|\mathcal{F}_{j-1}) \leq C$ a.s. for a positive constant $C > 0$.

It is of interest to extend the result of Theorem 1 to the stochastic regression models, where the regressors x_j are random variables. That stochastic regerssion model case deserves a further investigation and special attention.

REFERENCES

(1) Brockwell, P. J. and Davis, R. (1990). *Time Series: Theory and Methods, 2nd ed.* Springer, New York.

(2) Dunfour, J. M. (1986). Bias of s^2 in linear regression with dependent errors. *The American Statistician*, **40**, 284-285.

(3) Lai, T. L. and Wei, C. Z. (1982). Least squares estimates in stochastic regression model with applications to stochastic regression in linear dynamic systems. *Annals of Statistics*, **10**, 154-166.

(4) Neudecker, H. (1977). Bounds for the bias of the least squares estimator of σ^2 in the case of a first-order autoregressive process (positive autocorrelation). *Econometrica*, **45**, 1257-1262.

(5) Neudecker, H. (1978). Bounds for the LS estimator of σ^2 in the case of a first-order (positive) autoregressive process when the regression contains a constant term. *Econometrica*, **46**, 1223-1226.

(6) Sathe, S. T. and Vinod, H. D. (1974). Bounds on the variance of regression coefficients due to heteroscedastic or autoregressive errors. *Econometrica*, **42**, 333-340.

(7) Solo, V. (1981). Strong consistency of least squares estimators in regression with correlated disturbances. *Annals of Statistics*, **9**, 689-693.

(8) Song, H. S. (1994). The asymptotic unbiasedness of S^2 in the linear regression model with moving average or particular s-th order autocorrelated disturbances. *Journal of Korean Statistical Society*, **32**, 33-38.