ON THE EFFICIENCY OF THE SIMPLE LEAST SQUARES ESTIMATOR IN THE LINEAR MODEL UNDER $\mathcal{C}(V) \supset \mathcal{C}(X)$

Richard A. Stein, University of Arizona

1. INTRODUCTION. The linear model is often expressed as $y = X\beta + e$ where y is an $n \times 1$ vector of observations, X is an $n \times p$ matrix of known real constants, β is a $p \times 1$ vector of unknown parameters and e is an $n \times 1$ vector of errors with $E(e) = \phi$, the null vector, and with variance matrix $V = E(ee')$.

Let SLSE($X\beta$) denote the simple least squares estimator of $X\beta$ and BLUE($X\beta$) the minimum variance linear unbiased estimator of $X\beta$. This later can be expressed as $X(X'V^-X)^-X'V^-y$ whenever the range space of V contains the range space of X, Mitra & Rao (1968).

Durbin and Watson (1950) and Watson (1955) are among the first to lay ground work for comparing the performance of the simple least squares estimators with respect to the corresponding minimum variance linear unbiased estimators. Watson (1955) defined the efficiency of SLSE($X\beta$) to be $|X'X|^2/|X'VX||X'V^{-1}X|$. As this expression is the ratio of the generalized variances of SLSE($X\beta$) and BLUE($X\beta$) respectively when $X'X$ and V are invertible, it has considerable appeal. When $X'X$ or V are not invertible, certain difficulties arise.

Magness and McGuire (1962), considering $X'X$ and V invertible, establish relationships more along the lines of this paper, i.e. $\lambda_{min} \leq \mathrm{Var}[\mathrm{BLUE}(X\beta)] \leq \mathrm{Var}[\mathrm{SLSE}(X\beta)] \leq \lambda_{max}$ and $\mathrm{Var}[\mathrm{SLSE}(t'X\beta)] \leq (\lambda_{min}^{-1} + \lambda_{max}^{-1})(\lambda_{min} + \lambda_{max})\mathrm{Var}[\mathrm{BLUE}(t'X\beta)]/4$ where λ_{min} and λ_{max} are respectively the smallest and largest eigenvalues of V.

Golub (1963) extended the results of Magness and McGuire by using an inequality due to Schopf assuming invertibility of $X'X$ and V.

Rizzuto et. al. give an attainable efficiency defined as the same generalized variance ratio for the covariance structure $(1 - \rho)I + \rho J$. However, the bound of these authors appears to be attainable in the sense that for a particular covariance matrix there exists a design matrix for which their bound is attained.

Notationally, let A be any matrix. Then $\mathcal{C}(A)$ and $\mathcal{R}(A)$ are respectively the column and row spaces of A. Likewise, $\mathcal{C}^{\perp}(A)$ and $\mathcal{R}^{\perp}(A)$ are respectively the orthogonal complements of $\mathcal{C}(A)$ and $\mathcal{R}(A)$. The usual expectation and variance operators are $E(\cdot)$ and $\mathrm{Var}(\cdot)$. Finally, A^- will denote any generalized inverse of A and A^+ will be the Moore-Penrose pseudo-inverse of A.

2. GENERAL APPROACH. Let us define the efficiency as

$$\min_t \frac{\mathrm{Var}[\mathrm{BLUE}(t'X\beta)]}{\mathrm{Var}[\mathrm{SLSE}(t'X\beta)]} = \mathrm{Eff}(\mathrm{SLSE}(X\beta)) \qquad (1)$$

Two bounds, one being exact, will be given for this definition of efficiency under the constraint $\mathcal{C}(V) \supset \mathcal{C}(X)$. The meaning of exact is in the sense that the bound is attainable for whatever specific design matrix X is chosen.

Lemma 1: Eff(SLSE($X\beta$)) =

$$\frac{z'X'V^+Xz}{z'X'V^+X(X'X)^-X'VX(X'X)^-X'V^+Xz} \qquad (2)$$

where z is such that $t'X = z'X'V^+X$.

Such a z exists because $\mathcal{C}(V^+) = \mathcal{C}(V) \supset \mathcal{C}(X)$. Defining the denominator of (2) to be $f'f$, then

$$0 < 1/\max_{z'X'V^+Xz=1} f'f \leq \mathrm{Eff}(\mathrm{SLSE}(X\beta)) \leq 1/\min_{z'X'V^+Xz=1} f'f \leq 1$$

Via (2), the lower bound of the efficiency is seen to be attainable.

3. AN INEXACT LOWER BOUND.

THEOREM 1: In the model $y = X\beta + e$, $E(e) = \phi$, $\mathrm{Var}(e) = V$ with $\mathcal{C}(V) \supset \mathcal{C}(X)$,

$$\mathrm{Eff}(\mathrm{SLSE}(X\beta)) \geq \frac{\text{smallest nonzero scalar } h\text{: } X'VXt = hX'Xt}{\text{largest scalar } h\text{: } X'VXt = hX'Xt}$$

This bound is, in general, not attainable for any choice of X.

It can be shown that the largest eigenvalue of any symmetric positive semidefinite matrix A must, for all choices of X, exceed the largest scalar h such that $X'AXz = hX'Xz$. Also the smallest nonzero eigenvalue of A is smaller than or equal to the smallest h satisfying this expression. Thus via the use of h, the relationship, using appropriate normalized parametric functionals,

$$\lambda_{min} \leq h_{min} \leq \mathrm{Var}[\mathrm{BLUE}(t'X\beta)] \leq \mathrm{Var}[\mathrm{SLSE}(t'X\beta)] \leq h_{max} \leq \lambda_{max}$$

provides a tightening of one of the bounds of Magness and McGuire.

4. AN EXACT LOWER BOUND.

A second approach to efficiency bounds is to attempt to study directly $1/\text{Max } f'f$ subject to the constraint $z'X'V^+Xz = 1$. To this end, let the covariance matrix V be expressed in the form $V = (P|R|C)\begin{bmatrix} M & \phi & \phi \\ \phi & N & \phi \\ \phi & \phi & Q \end{bmatrix}\begin{pmatrix} P' \\ R' \\ C' \end{pmatrix}$ where M, N, Q are diagonal matrices. The columns of P form a complete set of orthonormal eigenvectors of V which lie in $\mathcal{C}(X)$. The columns of C form a complete set of orthonormal eigenvectors of V which lie in $\mathcal{C}^{\perp}(X)$. The columns of R are a complete set of the remaining orthonormal eigenvectors which are also orthogonal to $\mathcal{C}(P)$ and $\mathcal{C}(C)$.

All estimable parametric functionals can be expressed in the form $t'X\beta$. It is interesting that there exists a subspace of $\mathcal{R}(X)$ from which must come those $t'X$ for which $\text{Eff}(\text{SLSE}(X\beta))$ is attained.

THEOREM 2: In the model $y = X\beta + e$, $E(e) = \phi$, $E(ee') = V$, $\mathcal{C}(V) \supset \mathcal{C}(X)$, P, R, C as previously defined, $\text{Eff}(\text{SLSE}(X\beta))$ is attained for some parametric functional $t'X\beta$ if and only if $t'X \in \mathcal{C}(X'R)$.

The particular case where $\dim \mathcal{C}(X) - \dim \mathcal{C}(P) = 1$ has particular interest in that an attainable lower bound for $\text{Eff}(\text{SLSE}(X\beta))$ has a mathematical simplicity that seems otherwise lacking.

THEOREM 3: In the model $y = X\beta + e$, $E(e) = \phi$, $E(ee') = V$ with $\mathcal{C}(V) \supset \mathcal{C}(X)$, and $\dim \mathcal{C}(X) - \mathcal{C}(P) = 1$, let r be any vector in $\mathcal{C}(R)$ having that direction such that $\mathcal{C}(X) = \mathcal{C}(P) \oplus \mathcal{C}(r)$. Let $\{\cos \alpha_i, i = 1,\cdots,k\}$ be the set of directional cosines of the column vectors of R with respect to r and λ_i, the corresponding eigenvalues of V. Then $\text{Eff}(\text{SLSE}(X\beta)) = [(\sum_{i=1}^{k} \lambda_i^{-1} \cos^2\alpha_i)(\sum_{i=1}^{k} \lambda_i \cos^2\alpha_i]^{-1}$.

It should be noted that the smallest value of $\dim \mathcal{C}(R)$ is 2 whenever not all simple least squares estimators are best. This provides a setting for the following corollary.

COROLLARY: If under the conditions of Theorem 3, $\dim \mathcal{C}(R) = 2$, let $\cos \delta$ be the directional cosine of one of the two column vectors of R, and λ_{min}, λ_{max} are the two corresponding eigenvalues, then for fixed δ, $\text{Eff}(\text{SLSE}(X\beta))$ is strictly decreasing as $\lambda_{max} - \lambda_{min}$ increases. For fixed λ_{min}, λ_{max}, $\text{Eff}(\text{SLSE}(X\beta))$ is strictly decreasing as $|\delta - 45°|$ decreases.

THEOREM 4: In the model $y = X\beta + e$, $E(e) = \phi$, $E(ee') = V$, let $\mathcal{C}(V) \supset \mathcal{C}(X)$. Let P, R, C be as previously defined. Let X_2 have linearly independent column vectors such that $\mathcal{C}(X) = \mathcal{C}(P) \oplus \mathcal{C}(X_2)$ with $\dim \mathcal{C}(X_2) \geq 1$. Let ψ be any matrix where the i^{th} column vector is a set of directional cosines of the i^{th} column vector of X_2 with respect to the matrix $(P|R|C)$. Let W be any matrix satisfying $\mathcal{C}(W') = \mathcal{C}^{\perp}(\psi)$. Then $\text{Eff}(\text{SLSE}(X\beta)) = [(z'L^+z)(z'Lz)]^{-1}$ where L is diagonal such that $V = (P|R|C)L(P|R|C)'$ and z is any one of the solutions to $Wz = \phi$, $z'z = 1$,

$$[(z'L^+z)\psi'LL - 2(z'Lz)(z'L^+z)\psi'L + (z'Lz)\psi']z = \phi$$

COROLLARY: Under the conditions of Theorem 4, if the directional cosines of the matrix ψ are not considered, then $\text{Eff}(\text{SLSE}(X\beta)) \geq 4/(\lambda_{min}^{-1} + \lambda_{max}^{-1})(\lambda_{min} + \lambda_{max})$ where λ_{min} and λ_{max} are the smallest and largest eigenvalues of V with respect to the column vectors of R.

Here again the bounds for $\text{Eff}(\text{SLSE}(X\beta))$ of the preceding corollary are at least as close as those of Magness and McGuire since the maximum and minimum eigenvalues of the corollary are taken from a subset of all the eigenvalues of V.

5. REFERENCES

[1] Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression, I. Biometrika 37, 409-428.

[2] Golub, G.H. (1963). Comparison of the variances of the minimum variance and weighted least squares regression coefficients. Ann. Math. Stat. 34, 984-991.

[3] Magness, T.A. and McGuire, J.B. (1962). Comparison of least squares and minimum variance estimates of regression parameters. Ann. Math. Stat. 33, 462-470.

[4] Rizzuto, G.T., Boullion, T.L. and Lewis, T.O. (1973). An attainable lower bound for the efficiency of least squares estimates. 138th meeting IMS, Albuquerque.

[5] Watson, G.S. (1955). Serial correlation in regression analysis, I. Biometrika 42, 327-341.