English Translation

Chinese Journal of Applied Probability
and Statistics, Vol. 6, No. 3, Aug. 1990

Generalized Shrunken Least Squares Estimators*

Wang Liqun

Beijing Normal University

Abstract

This paper introduces a class of estimators for regression parameters in a linear model. This class contains many well-known linear estimators, such as ridge regression, principal components, shrunken least squares and the iteration estimators. The admissibility of the estimators in this class is studied and comparisons among them are made under the matrix mean square error criterion.

1 Introduction

Consider linear regression model

$$y = X\beta + \varepsilon, \tag{1.1}$$
$$E(\varepsilon) = 0,\ E(\varepsilon\varepsilon') = \sigma^2 I,$$

where X is a $T \times p$ known matrix, ε is a $T \times 1$ random vector, I is the identity matrix, $\beta \in \mathbb{R}^p$ and $0 < \sigma^2 < \infty$ are unknown parameters. Further, we assume that $2 \leq \operatorname{rank}(X) = p \leq T$.

It is well-known that under standard assumptions the ordinary least squares estimator

$$b_{OLS} = \left(X'X\right)^{-1} X'y$$

is the best linear unbiased or minimum variance unbiased estimator for β. However, as pointed out by Trenkler (1981), when any one of the assumptions of model (1.1) does not hold or the regression matrix X is ill-conditioned (e.g. there exists multicollinearity), b_{OLS} may perform very poorly. Consequently, in recent years many linear biased estimators have been developed to improve the performance of b_{OLS}. This paper considers the following class of linear estimators called the generalized shrunken least squares (GSLS) estimators:

$$b_{GS}(A) = PAP'b_{OLS},$$

where $A = \operatorname{diag}(a_1, a_2, ..., a_p)$, $0 \leq a_i \leq 1$, $i = 1, 2, ..., p$. Here P is the orthogonal matrix such that

$$P'X'XP = \operatorname{diag}(\lambda_1, \lambda_2, ..., \lambda_p) := \Lambda$$

and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$. It is easy to see that this class includes many common linear biased estimators. For example:

*Received February 6, 1985; revised October 25, 1988.

(1) The generalized ridge regression estimators of Hoerl and Kennard (1970):

$$b_{RR}(K) = \left(X'X + PKP'\right)^{-1} X'y = b_{GS}\left(\Lambda\left(\Lambda + K\right)^{-1}\right),$$

where $K = \operatorname{diag}(k_1, k_2, ..., k_p)$, $k_i \geq 0$, $i = 1, 2, ..., p$.

(2) The principal components estimators of Kendall (1957):

$$b_{PC}(r) = \sum_{i=1}^{r} \frac{1}{\lambda_i} P_i P_i' X'y = b_{GS}\left(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right),$$

where $P = (P_1, P_2, ..., P_p)$.

(3) The shrunken least squares estimators of Mayer and Willke (1973):

$$b_{SLS}(\alpha) = \alpha b_{OLS} = b_{GS}(\alpha I),$$

where $0 < \alpha < 1$.

(4) The iteration estimators of Trenkler (1978):

$$b_I(\alpha, n) = \alpha \sum_{i=0}^{n} \left(I - \alpha X'X\right)^i X'y = b_{GS}\left(I - (I - \alpha\Lambda)^{n+1}\right),$$

where $0 < \alpha < \lambda_1^{-1}$, $n = 0, 1, 2, ...$

Later we will see that, under the matrix mean square error criterion, in the class of generalized shrunken least squares estimators no one estimator is uniformly better than another one.

2 Admissibility of GSLS estimators

Let b_1 and b_2 be two estimators of β. Then b_1 is said to be better than b_2, if for all β and σ^2,

$$E\left[(b_1 - \beta)'(b_1 - \beta)\right] \leq E\left[(b_2 - \beta)'(b_2 - \beta)\right]$$

and the inequality holds for at least one pair (β, σ^2). In the following we denote the class of all homogeneous linear estimators of β as

$$\mathcal{L} = \{Ly \mid L \text{ is a } p \times T \text{ constant matrix}\}.$$

An estimator b of β is said to be $\mathcal{L}$–admissible, if $b \in \mathcal{L}$ and $\mathcal{L}$ does not contain any estimator which is better than b. By the result of Rao (1976) it is easy to see that every GSLS estimator $b_{GS}(A)$ is $\mathcal{L}$–admissible.

If an estimator b of β is admissible in the class of all estimators of β, then b is simply said to be admissible. To consider the admissibility of any GSLS estimator $b_{GS}(A)$ in the class of all estimators of β, in the rest of this section we assume that in model (1.1) it holds $\varepsilon \sim N\left(0, \sigma^2 I\right)$. We first prove a lemma, which is also of interest in its own right.

Lemma 2.1. *Assume random variable $y \sim N\left(\theta, \sigma^2 V\right)$, where $\theta \in I\!R^n$, $0 < \sigma^2 < \infty$ and V is a known positive definite matrix. Then for any linear estimator Ly, it is admissible for θ if and only if either 1) $L = I$; or 2) LV is symmetric, all eigenvalues of L are between 0 and 1 and at most two of them are equal to 1.*

Proof: First we show the sufficiency. For the case $L = I$, see Cheng (1982). For the case 2), we first consider a special case where $V = I$ and σ^2 is fixed. Denote $z = y/\sigma$. Then $z \sim N(\theta/\sigma, I)$. By the result of Cohen (1966), Lz is admissible for θ/σ. Therefore Ly is admissible for θ.

Next, consider the case where $y \sim N\left(\theta, \sigma^2 I\right)$ and σ^2 is unknown. We prove the result by contradiction. If there exists an estimator b of θ, such that for all θ and σ^2,

$$E\left[(b-\theta)'(b-\theta)\right] \leq E\left[(Ly-\theta)'(Ly-\theta)\right]$$

and the inequality holds for at least one pair (θ_0, σ_0^2), then for $y \sim N\left(\theta, \sigma_0^2 I\right)$, b will be better than Ly. However, this is impossible by the result already proved above.

Finally, consider the general $V > 0$. Let $z = V^{-1/2}y$, $\varphi = V^{-1/2}\theta$. Then $z \sim N\left(\varphi, \sigma^2 I\right)$. Under the lemma's conditions, $V^{-1/2}LV^{1/2}$ is symmetric, all its eigenvalues are between 0 and 1 and at most two of them are equal to 1. It follows that $V^{-1/2}LV^{1/2}z$ is admissible for φ and, therefore, Ly is admissible for θ.

Now we show the necessity. Suppose $L \neq I$. Since Ly is admissible for θ, it is also $\mathcal{L}$–admissible. By the result of Rao (1976), LV is symmetric and all eigenvalues of L are between 0 and 1. We need only to show that L has at most two eigenvalues being 1. Using the above notation, we have $z \sim N\left(\varphi, \sigma^2 I\right)$ and $V^{-1/2}LV^{1/2}z$ is admissible for φ. Hence without loss of generality we can assume $V = I$. Further, let U be an orthogonal matrix such that $U'LU = \operatorname{diag}(l_1, l_2, ..., l_n)$, where $1 \geq l_1 \geq l_2 \geq \cdots \geq l_n \geq 0$. Since $Uy \sim N\left(U\theta, \sigma^2 I\right)$, LUy is admissible for $U\theta$. It follows that $U'LUy$ is admissible for θ. Again we show the result by contradiction. If there exists $l_1 = l_2 = \cdots = l_r = 1$ and $r \geq 3$, then we construct an estimator which is better than $U'LUy$. Since $L \neq I$, we have $r \leq n-1$. Denote

$$y = (y_1, y_2, ..., y_n)', \ \theta = (\theta_1, \theta_2, ..., \theta_n)',$$

$$x = (y_1, y_2, ..., y_r)', \ \mu = (\theta_1, \theta_2, ..., \theta_r)'.$$

Then $x \sim N\left(\mu, \sigma^2 I\right)$.

If $r \leq n-2$, then let $S = \sum_{i=r+1}^{n}(y_i - \bar{y})^2$, where $\bar{y} = \sum_{i=r+1}^{n} y_i/(n-r)$. Otherwise if $r = n-1$, then let $S = y_n^2$. Then S and x are independent and $S/\sigma^2 \sim \chi^2(n-r)$. Further denote

$$b = (\omega y_1, \omega y_2, ..., \omega y_r, l_{r+1}y_{r+1}, l_{r+2}y_{r+2}, ..., l_n y_n)',$$

where

$$\omega = 1 - \frac{(r-2)S}{(n-r+2)x'x}.$$

Then we have

$$E\left[(b-\theta)'(b-\theta)\right] = E\left[(\omega x-\mu)'(\omega x-\mu)\right] + \sum_{i=r+1}^{n} E(l_i y_i - \theta_i)^2.$$

By the result of James and Stein (1961), for all θ and σ^2, it holds $E\left[(\omega x-\mu)'(\omega x-\mu)\right] < r\sigma^2$. Hence b will be uniformly better than $U'LUy$. The proof is completed. □

Now we have the main result of this section.

Theorem 2.1. *A GSLS estimator $b_{GS}(A)$ is admissible if and only if $\operatorname{rank}(I-A) \geq p-2$.*

Proof: An estimator $PAP'b_{OLS}$ is admissible for β if and only if $XPAP'b_{OLS}$ is admissible for $X\beta$. By Lemma 2.1 this is equivalent to that $XPAP'X'$ is symmetric, all eigenvalues of $XPA\Lambda^{-1}P'X'$

are between 0 and 1 and at most two of them are equal to 1. Since $XPA\Lambda^{-1}P'X'$ and A have the same eigenvalues, the necessary and sufficient condition is that A has at most two eigenvalues being 1, i.e., $\operatorname{rank}(I - A) \geq p - 2$. □

From this theorem we obtain the following results.

Corollary 2.1.

(1) b_{OLS} *is admissible if and only if* $p \leq 2$.

(2) $b_{RR}(K)$ *is admissible if and only if* $\operatorname{rank}(K) \geq p - 2$.

(3) $b_{PC}(r)$ *is admissible if and only if* $r \leq 2$.

(4) Any $b_{SLS}(\alpha)$ *is admissible.*

(5) Any $b_I(\alpha, n)$ *is admissible.*

Let $\mathcal{E}$ be the class of all estimators based on b_{OLS}. Under normality assumption we have $b_{OLS} \sim N\big(\beta, \sigma^2 (X'X)^{-1}\big)$. Completely analogously we can show the following results.

Theorem 2.2.

(1) b_{OLS} *is* $\mathcal{E}$*–admissible for* β.

(2) If $A \neq I$, *then* $PAP'b_{OLS}$ *is* $\mathcal{E}$*–admissible for* β *if and only if* $\operatorname{rank}(I - A) \geq p - 2$.

3 Comparison under matrix mean square error criterion

If b is an estimator of β, then the matrix mean square error of b is defined as

$$MSE(b) = E\left[(b - \beta)(b - \beta)'\right].$$

If $b_{GS}(A)$ is a GSLS estimator, then it is easy to see that

$$MSE(b_{GS}(A)) = \sigma^2 PA\Lambda^{-1}AP' + P(I - A)P'\beta\beta'P(I - A)P'.$$

Denote $\alpha = \frac{1}{\sigma^2}P'\beta$ and

$$\begin{aligned} M(A) &= \frac{1}{\sigma^2}P'\left[MSE(b_{GS}(A))\right]P \\ &= A\Lambda^{-1}A + (I - A)\alpha\alpha'(I - A). \end{aligned}$$

Then $b_{GS}(A)$ is better than $b_{GS}(B)$, i.e. $MSE(b_{GS}(A)) \leq MSE(b_{GS}(B))$ at point $\big(\beta, \sigma^2\big)$, if and only if

$$M(A) \leq M(B) \tag{3.1}$$

at $\alpha = P'\beta/\sigma^2$.

The main goal of this section is to compare any two GSLS estimators under the MSE criterion and to discuss some special cases. First we prove the following result.

Theorem 3.1. *For any two GSLS estimators* $b_{GS}(A)$ *and* $b_{GS}(B)$, *if* $M(A) \leq M(B)$ *for all* α, *then* $A = B$.

Proof: Under the theorem's condition it holds

$$\operatorname{trace}\left[\left(B^2 - A^2\right)\Lambda^{-1}\right] + \alpha'\left[(I-B)^2 - (I-A)^2\right]\alpha \geq 0$$

for all α. Since A and B are diagonal and satisfy $0 \leq A \leq I$ and $0 \leq B \leq I$, the above inequality implies $A = B$. □

Theorem 3.1 shows that in the class of GSLS estimators no one estimator is uniformly better than another one (i.e., for all β, σ^2). The following theorems provide conditions for (3.1) to hold. First we need the following lemma.

Lemma 3.1. *Let h and g be two $p \times 1$ vectors and let c be a constant. Then the eigenvalues of $H = hh' + cgg'$ are given by*

$$\frac{h'h + cg'g}{2} \pm \left[\frac{(h'h - cg'g)^2}{4} + c\left(h'g\right)^2\right]^{1/2}.$$

Proof: It is obvious that matrix H and

$$G = \begin{pmatrix} h' \\ g' \end{pmatrix} \begin{pmatrix} h & cg \end{pmatrix} = \begin{pmatrix} h'h & ch'g \\ g'h & cg'g \end{pmatrix}$$

have the same nonzero eigenvalues. It is also easy to calculate the characteristic equation for G as

$$\det(\lambda I - G) = \lambda^2 - \left(h'h + cg'g\right)\lambda + c\big(h'h \cdot g'g - (h'g)^2\big).$$

The conclusion follows. □

Theorem 3.2. *Let $b_{GS}(A)$ and $b_{GS}(B)$ be two GSLS estimators such that $A < B$ (i.e., $B - A$ is positive definite). Then $M(A) \leq M(B)$ if and only if*

$$u_1 - u_2 + u_1 u_2 - v_1^2 \leq 1, \tag{3.2}$$

where $u_1 = \alpha'\Lambda\left(B^2 - A^2\right)^+ (I-A)^2\alpha$, $u_2 = \alpha'\Lambda\left(B^2 - A^2\right)^+ (I-B)^2\alpha$ and $v_1 = \alpha'\Lambda\left(B^2 - A^2\right)^+ (I-A)(I-B)\alpha$. ($M^+$ denotes the Moore-Penrose inverse of matrix M.)

Proof: Denote $U = \left(B^2 - A^2\right)\Lambda^{-1}$ and $V = (I-A)\alpha\alpha'(I-A) - (I-B)\alpha\alpha'(I-B)$. Then

$$M(B) - M(A) = U - V = U^{1/2}\big(I - U^{-1/2}VU^{-1/2}\big)U^{1/2}.$$

Hence $M(A) \leq M(B)$ if and only if $U^{-1/2}VU^{-1/2} \leq I$, i.e., all eigenvalues of $U^{-1/2}VU^{-1/2}$ are less than or equal to 1. Now take $h = U^{-1/2}(I-A)\alpha$, $g = U^{-1/2}(I-B)\alpha$ and $c = -1$. Then it follows from Lemma 3.1 that the necessary and sufficient condition is

$$\frac{u_1 - u_2}{2} \pm \left[\frac{(u_1 + u_2)^2}{4} - v_1^2\right]^{1/2} \leq 1,$$

which is equivalent to (3.2). □

Theorem 3.3. *If $A \leq B$ and $A \neq B$, then $M(A) \leq M(B)$ if and only if (3.2) and*

$$(B-A)\alpha\alpha'(I-B)\left[I - (B-A)(B-A)^+\right] = 0 \tag{3.3}$$

hold.

Proof: Without loss of generality we assume

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

where $A_1 < B_1$. Correspondingly denote

$$\Lambda = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}.$$

Then

$$M(B) - M(A) = \begin{pmatrix} C_{11} & C_{12} \\ C_{12}' & 0 \end{pmatrix},$$

where

$$\begin{aligned} C_{11} &= \left(B_1^2 - A_1^2\right)\Lambda_1^{-1} + (I - B_1)\,\alpha_1\alpha_1'\,(I - B_1) - (I - A_1)\,\alpha_1\alpha_1'\,(I - A_1), \\ C_{12} &= (A_1 - B_1)\,\alpha_1\alpha_2'\,(I - A_2). \end{aligned}$$

By the property of non-negative definite matrices, $M(A) \le M(B)$ if and only if $C_{12} = 0$ and $C_{11} \ge 0$. By Theorem 3.2, $C_{11} \ge 0$ if and only if (3.2) holds for A_1, B_1, Λ_1 and α_1. However, under the theorem's condition, the left-hand side of (3.2) remains unchanged for A_1, B_1, Λ_1, α_1 and for A, B, Λ, α. Furthermore, it is easy to see that $C_{12} = 0$ is equivalent to (3.3). □

Theorem 3.4. *Let $b_{GS}(A)$ and $b_{GS}(B)$ be two GSLS estimators. If there are two or more i's such that $a_i > b_i$, then $M(A) \le M(B)$ can never hold.*

Proof: Without loss of generality suppose $a_i > b_i$ for $i = 1, 2$. Let A_1, B_1 and Λ_1 be the two dimensional matrices on the upper left corner of A, B and Λ respectively and let α_1 be the subvector of α consisting of the first two elements. Then

$$M(B) - M(A) = \begin{pmatrix} C_{11} & * \\ * & * \end{pmatrix},$$

where C_{11} is the same as in the proof of the previous theorem. Since

$$\operatorname{rank}\left[(I - B_1)\,\alpha_1\alpha_1'\,(I - B_1)\right] \le 1,$$

and $\left(B_1^2 - A_1^2\right)\Lambda_1^{-1} < 0$, the matrix $\left(B_1^2 - A_1^2\right)\Lambda_1^{-1} + (I - B_1)\,\alpha_1\alpha_1'\,(I - B_1)$ cannot be non-negative definite. It follows that $C_{11} \ge 0$ and, therefore, $M(A) \le M(B)$ cannot hold. □

In the following we need only consider the case where $a_i > b_i$ for at most one i. We consider the case

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & B_2 & 0 \\ 0 & 0 & A_3 \end{pmatrix},$$

where $b_1 < 1$, $A_2 < B_2$, b_1 is a constant and A_2, B_2, A_3 can be matrices. Correspondingly denote

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \Lambda_2 & 0 \\ 0 & 0 & \Lambda_3 \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Then

$$M(B) - M(A) = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ * & F_{22} & F_{23} \\ * & * & F_{33} \end{pmatrix},$$

where

$$\begin{aligned}
F_{11} &= \left(b_1^2-1\right)\lambda_1^{-1}+\left(1-b_1\right)^2\alpha_1^2, \\
F_{12} &= \left(1-b_1\right)\alpha_1\alpha_2'\left(I-B_2\right), \\
F_{13} &= \left(1-b_1\right)\alpha_1\alpha_3'\left(I-A_3\right), \\
F_{22} &= \left(B_2^2-A_2^2\right)\Lambda_2^{-1}+\left(I-B_2\right)\alpha_2\alpha_2'\left(I-B_2\right)-\left(I-A_2\right)\alpha_2\alpha_2'\left(I-A_2\right), \\
F_{23} &= \left(A_2-B_2\right)\alpha_2\alpha_3'\left(I-A_3\right).
\end{aligned}$$

Hence $M(A)\le M(B)$ if and only if $F_{13}=0$, $F_{23}=0$ and

$$\begin{pmatrix} F_{11} & F_{12} \\ * & F_{22} \end{pmatrix} \ge 0.$$

It is easy to see that the first two equalities above is equivalent to (3.3) and the matrix inequality holds if and only if

$$F_{11}=0, \quad F_{12}=0, \quad F_{22}\ge 0 \tag{3.4}$$

or

$$F_{11}>0, \quad F_{22}-F_{11}^{-1}F_{12}'F_{12}\ge 0. \tag{3.5}$$

Further denote

$$E_1=\begin{pmatrix} 0 & 0 \\ 0 & I_{p-1} \end{pmatrix}.$$

Then analogous to Theorem 3.2 it can be shown that (3.4) is equivalent to

$$\begin{cases} \alpha_1^2=\frac{1+b_1}{(1-b_1)\lambda_1}, \quad (I-B)(B-A)E_1\alpha=0, \\ u_1-u_3+u_1u_3-v_1^2\le 1, \end{cases} \tag{3.6}$$

where $u_3=\alpha'\Lambda\left(B^2-A^2\right)^+(I-B)^2E_1\alpha$, and u_1 and v_1 are the same as in (3.2). Further, (3.5) is equivalent to

$$\alpha_1^2>\frac{1+b_1}{\left(1-b_1\right)\lambda_1}, \qquad \frac{u_1+u_4}{2}+\left[\frac{\left(u_1-u_4\right)^2}{4}+v_2^2\right]^{1/2}\le 1, \tag{3.7}$$

where

$$u_4=\frac{\left(1+b_1\right)u_3}{\left(1-b_1\right)\lambda_1\alpha_1^2-\left(1+b_1\right)}, \qquad v_2=v_1\left[\frac{1+b_1}{\left(1-b_1\right)\lambda_1\alpha_1^2-1-b_1}\right]^{1/2}.$$

Thus we have shown the following result.

Theorem 3.5. *Suppose $b_{GS}(A)$ and $b_{GS}(B)$ are two GSLS estimators and A, B have the form*

$$A=\begin{pmatrix} 1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad B=\begin{pmatrix} b & 0 \\ 0 & B_2 \end{pmatrix},$$

where $b<1$ and $A_2\le B_2$. Then $M(A)\le M(B)$ if and only if either (3.3), (3.6) or (3.3), (3.7) hold.

The above theorems give the ranges of parameter values in which one GSLS estimator is better than another. Some expressions look rather complicated. If in practical applications only the sufficient or necessary condition is needed, then some simpler forms can be used. For example, a sufficient condition for (3.2) to hold is $u_1\le 1$.

By taking special forms of A and B, the above theorems give the comparison results among some common estimators (e.g., b_{OLS}, b_{RR}, b_{PC}, b_{SLS} and b_I). These results include that of Price (1982). To illustrate, we consider the following two special cases.

(i) Comparison between GSLS and b_{OLS}

Corollary 3.1. *Let $b_{GS}(A)$ be a GSLS estimator and $A \neq I$. Then*

(1) $M(A) \leq M(I)$ if and only if $\alpha'\Lambda(I-A)(I+A)^{-1}\alpha \leq 1$; and

(2) $M(I) \leq M(A)$ if and only if $\operatorname{rank}(I-A) \leq 1$ and $\alpha'\Lambda(I-A)(I+A)^{-1}\alpha \geq 1$.

(ii) Admissible principal components estimators

Corollary 3.2. *Suppose*

$$A = \begin{pmatrix} A_r & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix},$$

where $r \leq s < p$ and $A_r < I_r$. Then

(1) If $s \geq 2$, then $M(B) \leq M(A)$ can never hold; and

(2) $M(A) \leq M(B)$ if and only if $(I-B)\alpha = 0$ and $\alpha'\Lambda(B-A)(B+A)^{+}\alpha \leq 1$.

Acknowledgement: The author is grateful to Professor Wang Junxiang for his advice and helpful discussions during the preparation of this manuscript.

References

[1] Cheng Ping. Admissibility of simultaneous estimation of several parameters. *Journal of System Science and Mathematical Science*, **2**(1982), 176-195.

[2] Cohen, A. All admissible linear estimates of the mean vectors. *Ann. Math. Statist.*, **37**(1966), 458-463.

[3] Hoerl, A. E. and Kennard, R. W. Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics*, **12**(1970), 55-67.

[4] James, W. and Stein, C. Estimation with quadratic loss. *Proc. Fouth Berkeley Symp. Math. Statist. and Prob.*, **1**(1961), 361-379.

[5] Kendall, M. G. *A Course in Multivariate Analysis.* London: Charles Griffin, 1957.

[6] Mayer, L. S. and Willke, T. A. On biased estimation in linear models. *Technometrics*, **15**(1973), 497-508.

[7] Price, J. M. Comparisons among regression estimators under the generalized mean square error criterion. *Comm. Statist.*, **AII**(1982), 1965-1984.

[8] Rao, C. R. Estimation of parameters in a linear model. *Ann. Math. Statist.*, **4**(1976), 1023-1037.

[9] Trenkler, G. An iteration estimator for the linear model. *COMPSTAT 1978*, Wien, 125-131. Physica Verlag, Würzburg, 1978.

[10] Trenkler, G. *Biased Estimators in the Linear Regression Model.* Oelgeschlager, Gunn and Hain, 1981.