Scientiae Mathematicae Japonicae Online, **Vol. 4**, (2001), 65–74

A NOTE ON AN UPPER BOUND FOR THE COVARIANCE MATRIX OF A GENERALIZED LEAST SQUARES ESTIMATOR IN A HETEROSCEDASTIC MODEL

Hiroshi Kurata

Received March 10, 2000; revised June 16, 2000

Abstract. This paper complements the results of Toyooka and Kariya (*The Annals of Statistics, 1986*) by evaluating an upper bound for the covariance matrix of a typical generalized least squares estimator in a heteroscedastic model.

1 Introduction. We consider the problem of estimating the coefficient vector of a heteroscedastic model with two distinct variances. We investigate the efficiency of a typical generalized least squares estimator (GLSE) by evaluating an upper bound for its covariance matrix. The upper bound considered here is the one that Toyooka and Kariya (1986) derived under a more general setup. In Toyooka and Kariya (1986), under a general linear regression model with a certain covariance structure which includes a serial correlation model and a heteroscedastic model as its special cases, the general formula of an upper bound for the risk matrix of a GLSE was obtained. The result was further applied to several typical GLSE's in the above two specific models. However, as for the heteroscedastic model, their evaluation was not explicit and was limited to the common mean model, a special case of the heteroscedastic model. In this paper, we complement their result by treating an unbiased GLSE based on the ratio of the two sample variances and obtaining an explicit expression of the upper bound for its covariance matrix.

In Section 2, we briefly review the results of Toyooka and Kariya (1986) on which our discussion is based. In Section 3, we present the main result. In Section 4, we investigate the relation between the upper bound evaluated in Section 3 and an alternative upper bound which has been considered from a different point of view by Kariya (1981), Bilodeau (1990) and Kurata and Kariya (1996).

2 Preliminaries. The heteroscedastic model considered here is given by

$$\boldsymbol{y}_j = \boldsymbol{X}_j\boldsymbol{\beta} + \boldsymbol{\varepsilon}_j \quad (j = 1, 2), \tag{1}$$

where $\boldsymbol{y}_j : n_j \times 1$, $\boldsymbol{X}_j : n_j \times k$, $rank\boldsymbol{X}_j = k$, $\boldsymbol{\beta} : k \times 1$, $\boldsymbol{\varepsilon}_j : n_j \times 1$ and the error terms $\boldsymbol{\varepsilon}_j$'s are supposed to be independently distributed as the normal distribution $N_{n_j}(\boldsymbol{0}, \sigma_j^2 \boldsymbol{I}_{n_j})$. The model (1) is a special case of the following general linear regression model of the form

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad with \quad E(\boldsymbol{\varepsilon}) = \boldsymbol{0} \;\; and \;\; Cov(\boldsymbol{\varepsilon}) = \boldsymbol{\Omega} = \sigma^2\boldsymbol{\Sigma}, \tag{2}$$

where $\boldsymbol{X}$ is an $n \times k$ known matrix of rank k, $\boldsymbol{\Omega} = \sigma^2\boldsymbol{\Sigma}$ is positive definite and $\boldsymbol{\Sigma}$ is a function of an unknown but estimable parameter θ, say $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\theta)$. In fact, letting $n = n_1 + n_2$,

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix}, \; \boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \; \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \end{pmatrix} \tag{3}$$

2000 *Mathematics Subject Classification.* 62J05, 62H12 .

Key words and phrases. Generalized least squares estimator, Heteroscedastic model, Upper bound for covariance matrix .

and $\boldsymbol{\Omega} = \sigma^2 \boldsymbol{\Sigma}$ with $\sigma^2 = \sigma_1^2$, $\theta = \sigma_1^2/\sigma_2^2$ and

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\theta) = \begin{pmatrix} \boldsymbol{I}_{n_1} & \boldsymbol{0} \\ \boldsymbol{0} & \theta^{-1}\boldsymbol{I}_{n_2} \end{pmatrix} \tag{4}$$

in (2) obviously yields the model (1).

Under the general model (2), the Gauss-Markov estimator (GME)

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma}) = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y}$$

is the best linear unbiased estimator of $\boldsymbol{\beta}$, provided that θ in $\boldsymbol{\Sigma}(\theta)$ is known. In our case where θ is unknown, a GLSE of the form

$$\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}}) = (\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y} \quad \textit{with} \quad \hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\hat{\theta}) \tag{5}$$

is often used, where $\hat{\theta} = \hat{\theta}(\boldsymbol{e})$ is an estimator of θ based on the ordinary least squares (OLS) residual vector

$$\boldsymbol{e} = \boldsymbol{N}\boldsymbol{y} \quad \textit{with} \quad \boldsymbol{N} = \boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'. \tag{6}$$

Kariya (1985), Kariya and Toyooka (1985) and Eaton (1985) proved that if the distribution of ε satisfies

$$E(\boldsymbol{u}_1|\boldsymbol{u}_2) = \boldsymbol{0} \quad a.s., \tag{7}$$

then for any GLSE of the form (5), its risk matrix is bounded below by the covariance matrix of the GME, that is,

$$Cov\left(\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma})\right) \le E\left((\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}}) - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}}) - \boldsymbol{\beta})'\right) \tag{8}$$

holds, where $\boldsymbol{u}_1 = \boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\varepsilon$, $\boldsymbol{u}_2 = \boldsymbol{Z}'\varepsilon$ and $\boldsymbol{Z}$ is any $n \times (n-k)$ matrix such that $\boldsymbol{X}'\boldsymbol{Z} = \boldsymbol{0}$ and $\boldsymbol{Z}'\boldsymbol{Z} = \boldsymbol{I}_{n-k}$. We note that the condition (7) is satisfied when ε is normally distributed.

Based on the inequality (8), Toyooka and Kariya (1986) derived an upper bound for the risk matrix of a GLSE relative to the covariance matrix of the GME in the case where $\boldsymbol{\Sigma}$ in (2) has the following structure

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}(\theta)^{-1} = \boldsymbol{I}_n + \lambda(\theta)\boldsymbol{D} \quad \textit{with} \quad \theta \in \Theta. \tag{9}$$

Here Θ is an open and nonempty interval in R^1, $\lambda(\theta)$ is a continuous function on Θ, and $\boldsymbol{D}$ is a known nonnegative definite matrix. As is discussed in their paper, the heteroscedastic model treated here satisfies the condition (9), since the matrix $\boldsymbol{\Sigma}$ in (4) is rewritten as (9) by letting $\Theta = (0, \infty)$, $\lambda(\theta) = \theta - 1$ and

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n_2} \end{pmatrix}.$$

In the following lemma, the upper bound obtained in Toyooka and Kariya (1986) is presented in the context of the heteroscedastic model. (See also Section 3 of their paper.)

Lemma 1 (Toyooka and Kariya (1986)) *In the heteroscedastic model (3) with (4), suppose that an estimator $\hat{\theta}$ satisfies $\hat{\theta} \in \Theta$ a.s., and let*

$$B_1 = \{\, \boldsymbol{y} \in R^n \mid \hat{\theta} \ge \theta \}, \quad B_2 = \{\, \boldsymbol{y} \in R^n \mid \hat{\theta} < \theta \}, \quad W_1 = 1, \quad W_2 = \theta^2/\hat{\theta}^2, \quad \boldsymbol{F} = \theta^{-1}\boldsymbol{D}$$

and

$$\boldsymbol{M} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1/2}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1/2}\boldsymbol{F}\boldsymbol{\Sigma}^{-1/2}\left[\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\right]. \tag{10}$$

Then for the GLSE $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}})$ *with* $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\hat{\theta})$*,*

$$E\left((\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}}) - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}}) - \boldsymbol{\beta})'\right) \le [1 + h_1 + h_2]Cov\left(\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma})\right) \tag{11}$$

holds, where $h_j = E(H_j)$*,*

$$H_j = \mathbf{1}_{B_j}(\hat{\theta} - \theta)^2 W_j \varepsilon'\boldsymbol{M}'\boldsymbol{M}\varepsilon/\sigma^2 \quad (j = 1, 2) \tag{12}$$

and $\mathbf{1}_B$ *denotes the indicator function of a set* B*.*

In Toyooka and Kariya (1986), the evaluation of h_j was limited to the common mean model which is obtained by letting $k = 1$ and $\boldsymbol{X} = (1, \cdots, 1)' : n \times 1$ in the model (3). Further, they did not derive its explicit expression. In Section 3 and 4, we derive it under the model (3) with general $\boldsymbol{X}$, and clarify the structure of the upper bound.

3 Evaluation of Upper Bound. For the heteroscedastic model (3) with (4), a minimal sufficient statistic is given by $(\boldsymbol{b}_1, \boldsymbol{b}_2, s_1^2, s_2^2)$, where

$$\boldsymbol{b}_j = (\boldsymbol{X}_j'\boldsymbol{X}_j)^{-1}\boldsymbol{X}_j'\boldsymbol{y}_j \tag{13}$$

and

$$s_j^2 = \boldsymbol{y}_j'\boldsymbol{N}_j\boldsymbol{y}_j = \varepsilon_j'\boldsymbol{N}_j\varepsilon_j \quad \textit{with} \quad \boldsymbol{N}_j = \boldsymbol{I}_{n_j} - \boldsymbol{X}_j(\boldsymbol{X}_j'\boldsymbol{X}_j)^{-1}\boldsymbol{X}_j'. \tag{14}$$

The statistics $\boldsymbol{b}_j$'s and s_j^2/σ_j^2's are independently distributed as $N_k(\boldsymbol{\beta}, \sigma_j^2(\boldsymbol{X}_j'\boldsymbol{X}_j)^{-1})$ and $\chi^2_{m_j}$, the χ^2-distribution with degrees of freedom $m_j = n_j - k$, respectively.

A typical estimator of θ is of the form

$$\hat{\theta} = cs_1^2/s_2^2 \quad \textit{with} \quad c > 0. \tag{15}$$

To apply Lemma 1, we confirm that the estimator $\hat{\theta}$ is a function of the OLS residual vector $\boldsymbol{e}$ in (6). Letting $\boldsymbol{e} = (\boldsymbol{e}_1', \boldsymbol{e}_2')'$ such that $\boldsymbol{e}_j : n_j \times 1$, we decompose $\boldsymbol{e}_1$ into two independent parts as

$$\begin{aligned}
\boldsymbol{e}_1 &= \boldsymbol{y}_1 - \boldsymbol{X}_1(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \\
&= [\boldsymbol{X}_1\boldsymbol{b}_1 + \boldsymbol{N}_1\varepsilon_1] - \boldsymbol{X}_1(\boldsymbol{X}'\boldsymbol{X})^{-1}\left[\boldsymbol{X}'\boldsymbol{X}\boldsymbol{b}_1 + \boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{b}_2 - \boldsymbol{b}_1)\right] \\
&= \boldsymbol{N}_1\varepsilon_1 - \boldsymbol{X}_1(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{b}_2 - \boldsymbol{b}_1),
\end{aligned}$$

from which $s_1^2 = \boldsymbol{e}_1'\boldsymbol{N}_1\boldsymbol{e}_1$ follows. Similarly, we obtain $s_2^2 = \boldsymbol{e}_2'\boldsymbol{N}_2\boldsymbol{e}_2$ and thus we can apply Lemma 1 to the GLSE of the form

$$\begin{aligned}
\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}}) &= (\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y} \quad \textit{with} \;\; \hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\hat{\theta}) \\
&= (\boldsymbol{X}_1'\boldsymbol{X}_1 + \hat{\theta}\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}(\boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{b}_1 + \hat{\theta}\boldsymbol{X}_2'\boldsymbol{X}_2\boldsymbol{b}_2).
\end{aligned} \tag{16}$$

As is well known, the estimator $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\Sigma}})$ is unbiased and has finite second moments for all $c > 0$. A typical choice of c will be $c = m_2/m_1$. A lot of researches have been made on the efficiency of the GLSE in the literature. See, for example, Khatri and Shah (1974), Taylor (1977,1978), Swamy and Mehta (1979) and Kubokawa (1998) in addition to the papers given in the previous sections.

Lemma 2 *For any GLSE of the form (16), the quantity $H_1 + H_2$ in (12) is expressed as*

$$H_1 + H_2 = \ell(\hat{\theta}, \theta)\ Q(\boldsymbol{b}_1 - \boldsymbol{b}_2), \tag{17}$$

where the functions ℓ and Q are defined as

$$\ell(\hat{\theta}, \theta) = \ell_0(\hat{\theta}/\theta) + \ell_0(\theta/\hat{\theta}) \quad with \quad \ell_0(t) = (t-1)^2 \mathbf{1}_{\{t>1\}} \tag{18}$$

and

$$Q(\boldsymbol{x}) = \frac{\theta^2}{\sigma^2}\ \boldsymbol{x}'\boldsymbol{V}\boldsymbol{x}$$

with

$$\boldsymbol{V} = \boldsymbol{X}_1'\boldsymbol{X}_1(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}_1'\boldsymbol{X}_1, \tag{19}$$

respectively.

Proof. From (12), we readily obtain $H_1 = \ell_0(\hat{\theta}/\theta)\ \boldsymbol{\varepsilon}'\boldsymbol{M}'\boldsymbol{M}\boldsymbol{\varepsilon}\ (\theta^2/\sigma^2)$ and $H_2 = \ell_0(\theta/\hat{\theta})\ \boldsymbol{\varepsilon}'\boldsymbol{M}'\boldsymbol{M}\boldsymbol{\varepsilon}\ (\theta^2/\sigma^2)$. Hence we have

$$H_1 + H_2 = \ell(\hat{\theta}, \theta)\ \boldsymbol{\varepsilon}'\boldsymbol{M}'\boldsymbol{M}\boldsymbol{\varepsilon}\ (\theta^2/\sigma^2).$$

Since $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{F}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{D}$, we get

$$\boldsymbol{M}\boldsymbol{\varepsilon} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1/2}\boldsymbol{X}'\boldsymbol{D}\left[\boldsymbol{I}_n - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\right]\boldsymbol{\varepsilon}. \tag{20}$$

Substituting

$$\boldsymbol{X}'\boldsymbol{D}\boldsymbol{\varepsilon} = \boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{b}_2 - \boldsymbol{\beta})$$

and

$$\begin{aligned} &\boldsymbol{X}'\boldsymbol{D}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} \\ &= \boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\left[\boldsymbol{X}_1'\boldsymbol{X}_1(\boldsymbol{b}_1 - \boldsymbol{b}_2) + (\boldsymbol{X}_1'\boldsymbol{X}_1 + \theta\boldsymbol{X}_2'\boldsymbol{X}_2)(\boldsymbol{b}_2 - \boldsymbol{\beta})\right] \\ &= \boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}_1'\boldsymbol{X}_1(\boldsymbol{b}_1 - \boldsymbol{b}_2) + \boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{b}_2 - \boldsymbol{\beta}) \end{aligned} \tag{21}$$

into (20) proves (17), where the last equality of (21) is due to

$$\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} = \boldsymbol{X}_1'\boldsymbol{X}_1 + \theta\boldsymbol{X}_2'\boldsymbol{X}_2. \square$$

Since $\hat{\theta}$ depends only on s_j^2's, the quantities $\ell(\hat{\theta}, \theta)$ and $Q(\boldsymbol{b}_1 - \boldsymbol{b}_2)$ are independent. Hence we see that

$$\begin{aligned} h_1 + h_2 &= E(H_1 + H_2) = E\left(\ell(\hat{\theta}, \theta)\right) E\left(Q(\boldsymbol{b}_1 - \boldsymbol{b}_2)\right) \\ &\equiv \rho(c; m_1, m_2)\ E\left(Q(\boldsymbol{b}_1 - \boldsymbol{b}_2)\right), \end{aligned} \tag{22}$$

where the last equality defines the function $\rho(c; m_1, m_2) = E(\ell(\hat{\theta}, \theta))$. The second factor of (22) is further calculated in the following lemma.

Lemma 3

$$E\left(Q(\boldsymbol{b}_1 - \boldsymbol{b}_2)\right) = \sum_{i=1}^{k} \frac{\theta r_i}{(1 + \theta r_i)^2} \equiv q(r_1, \cdots, r_k; \theta) \quad (say), \tag{23}$$

where $r_1, \cdots, r_k$ are the latent roots of the matrix $(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1/2}\boldsymbol{X}_2'\boldsymbol{X}_2(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1/2}$.

Proof. Let $\boldsymbol{G}$ be a $k \times k$ nonsingular matrix such that

$$\boldsymbol{X}_1'\boldsymbol{X}_1 = \boldsymbol{G}\boldsymbol{G}' \;\; and \;\; \boldsymbol{X}_2'\boldsymbol{X}_2 = \boldsymbol{G}\boldsymbol{R}\boldsymbol{G}' \quad with \quad \boldsymbol{R} = \boldsymbol{diag}\{r_1, \cdots, r_k\},$$

where $\boldsymbol{diag}$ denotes the diagonal matrix. Then the matrices $\boldsymbol{V}$ in (19) and $Cov(\boldsymbol{b}_1 - \boldsymbol{b}_2)$ are expressed as

$$\begin{aligned} \boldsymbol{V} &= \boldsymbol{G}\boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}' + \theta \boldsymbol{G}\boldsymbol{R}\boldsymbol{G}')^{-1}\boldsymbol{G}\boldsymbol{R}\boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}' + \theta \boldsymbol{G}\boldsymbol{R}\boldsymbol{G}')^{-1}\boldsymbol{G}\boldsymbol{R}\boldsymbol{G}'(\boldsymbol{G}\boldsymbol{G}' + \theta \boldsymbol{G}\boldsymbol{R}\boldsymbol{G}')^{-1}\boldsymbol{G}\boldsymbol{G}' \\ &= \boldsymbol{G}\ \boldsymbol{diag}\{r_1^2/(1+\theta r_1)^3, \cdots, r_k^2/(1+\theta r_k)^3\}\ \boldsymbol{G}' \end{aligned}$$

and

$$\begin{aligned} Cov(\boldsymbol{b}_1 - \boldsymbol{b}_2) &= \sigma^2[(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1} + \theta^{-1}(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}] \\ &= \sigma^2 \boldsymbol{G}'^{-1}\ \boldsymbol{diag}\{(1+\theta r_1)/(\theta r_1), \cdots, (1+\theta r_k)/(\theta r_k)\}\ \boldsymbol{G}^{-1}, \end{aligned}$$

respectively. Thus the result follows since $E(Q(\boldsymbol{b}_1 - \boldsymbol{b}_2)) = (\theta^2/\sigma^2)tr(\boldsymbol{V}\ Cov(\boldsymbol{b}_1 - \boldsymbol{b}_2))$.□
Note that the function q in (23) does not depend on the choice of c in $\hat{\theta} = cs_1^2/s_2^2$.

On the other hand, the quantity $\rho(c; m_1, m_2) = E(\ell(\hat{\theta}, \theta))$ reflects the loss of efficiency caused by estimating unknown θ. As a loss function for estimating θ, the function $\ell(\hat{\theta}, \theta)$ depends only on $\hat{\theta}/\theta$, say $\ell(\hat{\theta}, \theta) = \tilde{\ell}(\hat{\theta}/\theta)$. It has the following symmetric inverse property

$$\tilde{\ell}(t) = \tilde{\ell}(1/t) \quad for\ any\ t > 0,$$

which means that the loss function equally penalizes the underestimate and the overestimate of θ. The term "symmmetric inverse" is due to Bilodeau (1990).

To describe $\rho(c; m_1, m_2)$, we use the hypergeometric function

$${}_2F_1(a_1, a_2; a_3; z) = \sum_{j=0}^{\infty} \frac{(a_1)_j (a_2)_j}{(a_3)_j} \frac{z^j}{j!} \quad with \;\; (a)_j = \prod_{i=0}^{j-1}(a+i) \;\; and \;\; (a)_0 = 1,$$

which converges for $|z| < 1$. (See, for example, Abramowitz and Stegun (1972))

Theorem 1 *Let $m_j > 4$ $(j = 1, 2)$. Then, for any $\hat{\theta} = cs_1^2/s_2^2$ $(c > 0)$, the equality*

$$\rho(c; m_1, m_2) = \begin{cases} \rho_1(c; m_1, m_2) & (0 < c < 1) \\ \rho_2(m_1, m_2) & (c = 1) \\ \rho_3(c; m_1, m_2) & (1 < c) \end{cases}$$

holds, where ρ_1, ρ_2 and ρ_3 are given by

(24)

$$\begin{aligned} \rho_1(c; m_1, m_2) = \frac{1}{B\left(\frac{m_1}{2}, \frac{m_2}{2}\right)} \Bigg[& B\left(\frac{m_2}{2} - 2, 3\right) c^{m_2/2} {}_2F_1\left(\frac{m_1+m_2}{2}, \frac{m_2}{2} - 2; \frac{m_2}{2} + 1;\ -c\right) \\ & + B\left(\frac{m_1}{2} - 2, 3\right) c^{m_2/2} \left(\frac{1}{1+c}\right)^{(m_1+m_2)/2} {}_2F_1\left(\frac{m_1+m_2}{2}, 3; \frac{m_1}{2} + 1; \frac{1}{1+c}\right) \Bigg], \end{aligned}$$

(25)

$$\begin{aligned} \rho_2(m_1, m_2) \;\; = \;\; & \frac{1}{2^{(m_1+m_2)/2} B\left(\frac{m_1}{2}, \frac{m_2}{2}\right)} \left[B\left(\frac{m_2}{2} - 2, 3\right) {}_2F_1\left(\frac{m_1+m_2}{2}, 3; \frac{m_2}{2} + 1; \frac{1}{2}\right) \right. \\ & \left. + B\left(\frac{m_1}{2} - 2, 3\right) {}_2F_1\left(\frac{m_1+m_2}{2}, 3; \frac{m_1}{2} + 1; \frac{1}{2}\right) \right], \end{aligned}$$

and

$$\rho_3(c; m_1, m_2) = \rho_1(1/c; m_2, m_1), \tag{26}$$

respectively. (Thus, for any GLSE $\hat{\boldsymbol\beta}(\boldsymbol\Sigma(\hat\theta))$ *with* $\hat\theta = cs_1^2/s_2^2$*,*

$$Cov\left(\hat{\boldsymbol\beta}(\boldsymbol\Sigma(\hat\theta))\right) \le [1 + \rho(c; m_1, m_2)q(r_1, \cdots, r_k; \theta)]\, Cov\left(\hat{\boldsymbol\beta}(\boldsymbol\Sigma(\theta))\right)$$

holds.)

Proof. Let $v_j = s_j^2/\sigma_j^2$ so that $v_j \sim \chi^2_{m_j}$. Then, by (18), we have

$$\rho(c; m_1, m_2) = E\left(\ell_0(cv_1/v_2)\right) + E\left(\ell_0(v_2/cv_1)\right), \tag{27}$$

from which the following equality

$$\rho(c; m_1, m_2) = \rho(1/c; m_2, m_1) \tag{28}$$

is obtained.

We first prove (24) and (25). Letting $a = 1/(2^{(m_1+m_2)/2}\Gamma(m_1/2)\Gamma(m_2/2))$, the first term of the right hand side of (27) is written as

$$E\left(\ell_0(cv_1/v_2)\right) = a \int\int_{cv_1/v_2 \ge 1} (cv_1/v_2 - 1)^2 v_1^{m_1/2-1} v_2^{m_2/2-1} \exp\left(-(v_1+v_2)/2\right) dv_1 dv_2.$$

Making transformation $z_1 = v_1$ and $z_2 = v_2/(cv_1)$ with $dv_1dv_2 = cz_1dz_1dz_2$ and integrating with respect to z_1 yields

$$\begin{aligned} & a\, c^{m_2/2} \int_0^1 z_2^{(m_2/2-2)-1}(1-z_2)^2 \left(\int_0^\infty z_1^{(m_1+m_2)/2-1} \exp\left(-(1+cz_2)z_1/2\right) dz_1\right) dz_2 \\ = \;& a' \int_0^1 z_2^{(m_2/2-2)-1}(1-z_2)^2(1+cz_2)^{-(m_1+m_2)/2} dz_2 \end{aligned} \tag{29}$$

with $a' = a \times c^{m_2/2}\Gamma((m_1+m_2)/2)2^{(m_1+m_2)/2} = c^{m_2/2}/B(m_1/2, m_2/2)$. To evaluate (29) in the case where $0 < c < 1$, we use the following well known formula

$$\int_0^1 t^{a_2-1}(1-t)^{a_3-a_2-1}(1-zt)^{-a_1} dt = B(a_2, a_3 - a_2)\; {}_2F_1(a_1, a_2; a_3; z), \tag{30}$$

which is valid for $0 < a_2 < a_3$ and $|z| < 1$. Applying the formula (30) to the right hand side of (29) proves the first term of the right hand side of (24). When $1 \le c$, applying the following formula

(31)
$$\int_0^1 t^{a_2-1}(1-t)^{a_3-a_2-1}(1-zt)^{-a_1} dt = (1-z)^{-a_1} B(a_2, a_3 - a_2)\; {}_2F_1\left(a_1, a_3 - a_2; a_3; \frac{z}{z-1}\right),$$

which is valid for $0 < a_2 < a_3$ and $z \le -1$, establishes

(32)
$$E\left(\ell_0(cv_1/v_2)\right) = \frac{B(m_2/2-2, 3)}{B(m_1/2, m_2/2)}(1+c)^{-(m_1+m_2)/2} c^{m_2/2}\, {}_2F_1\left(\frac{m_1+m_2}{2}, 3; \frac{m_2}{2}+1; \frac{c}{1+c}\right).$$

Substituting $c = 1$ into (32) yields the first term of (25).

Next we consider the second term of the right hand side of (27). By interchanging m_1 and m_2 and replacing c by $1/c$ in (29), we obtain

$$(33) \qquad E\left(\ell_0(v_2/cv_1)\right) \quad = \quad a'' \int_0^1 z_1^{(m_1/2-2)-1}(1-z_1)^2(1+c^{-1}z_1)^{-(m_1+m_2)/2}dz_1$$

with $a'' = 1/(c^{m_1/2}B(m_1/2, m_2/2))$. Since $0 < c \leq 1$ is equivalent to $1 \leq c^{-1}$, applying (31) to the right hand side of (33) establishes the second term of (24). The second term of (25) is obtained by letting $c = 1$. Thus (24) and (25) are proved. Finally, (26) is obtained from (28). $\square$

In the table below, we treat the GLSE with $c = m_2/m_1$ and summarize the values of $\rho(m_2/m_1; m_1, m_2)$ for $m_1, m_2 = 5, 10, 15, 20, 25, 50$. The table is symmetric in m_1 and m_2, which is a consequence of (28). We can observe that the upper bound monotonically decreases in m_1 and m_2.

Table of $\rho(m_2/m_1; m_1, m_2)$

m_2 \ m_1	5	10	15	20	25	50
5	18.3875	8.8474	7.7853	7.3423	7.0956	6.6356
10	8.8474	1.8203	1.2826	1.0879	0.9873	0.8138
15	7.7853	1.2826	0.8156	0.6518	0.5687	0.4288
20	7.3423	1.0879	0.6518	0.5013	0.4258	0.3001
25	7.0956	0.9873	0.5687	0.4258	0.3544	0.2369
50	6.6356	0.8138	0.4288	0.3001	0.2369	0.1349

4 Comparison with Another Upper Bound. In this section, we investigate the relation between the upper bound $1 + \rho(c; m_1, m_2)\ q(r_1, \cdots, r_k; \theta)$ and an alternative upper bound considered in the literature. In Kariya (1981) and Bilodeau (1990), it is shown that for any GLSE $\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma}(\hat{\theta}))$ of the form (16), the following inequality

$$(34) \qquad Cov\left(\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma}(\hat{\theta}))\right) \leq \left[1 + E(a(\hat{\theta}, \theta))\right] Cov\left(\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma}(\theta))\right)$$

holds, where $a(\hat{\theta}, \theta)$ is a symmetric inverse loss function given by

$$a(\hat{\theta}, \theta) = \frac{1}{4}\left[\hat{\theta}/\theta + \theta/\hat{\theta} - 2\right] = \frac{(\hat{\theta}/\theta - 1)^2}{4\hat{\theta}/\theta}.$$

The inequality (34) is derived from the structure of the conditional covariance matrix of $\hat{\boldsymbol{\beta}}(\boldsymbol{\Sigma}(\hat{\theta}))$ given $\hat{\theta}$. An extension of this result is given in Kurata and Kariya (1996) and Kurata (1999).

We show that the relation between the two upper bounds is indefinite. More precisely,

Theorem 2 (i) *The relation between $\ell(\hat{\theta}, \theta)$ and $a(\hat{\theta}, \theta)$ is given by*

$$(35) \qquad \ell(\hat{\theta}, \theta)/4 \geq a(\hat{\theta}, \theta).$$

(ii) *The range of the function q is given by*

$$0 < q(r_1, \cdots, r_k; \theta) \leq k/4,$$

and its maximum is attained when $r_1 = \cdots = r_k = 1/\theta$. As $r_i\theta$ goes to either 0 or ∞ $(i = 1, \cdots, k)$, the function q converges to 0.

Proof. (i) is proved as

$$
\begin{aligned}
\ell(\hat{\theta},\theta) &= (\hat{\theta}/\theta-1)^2\mathbf{1}_{\{\hat{\theta}/\theta>1\}}+(\theta/\hat{\theta}-1)^2\mathbf{1}_{\{\theta/\hat{\theta}>1\}} \\
&\geq \frac{(\hat{\theta}/\theta-1)^2}{\hat{\theta}/\theta}\mathbf{1}_{\{\hat{\theta}/\theta>1\}}+\frac{(\theta/\hat{\theta}-1)^2}{\theta/\hat{\theta}}\mathbf{1}_{\{\theta/\hat{\theta}>1\}} \\
&= \frac{(\hat{\theta}/\theta-1)^2}{\hat{\theta}/\theta}\mathbf{1}_{\{\hat{\theta}/\theta>1\}}+\frac{(\hat{\theta}/\theta-1)^2}{\hat{\theta}/\theta}\mathbf{1}_{\{\hat{\theta}/\theta<1\}} \\
&= \frac{(\hat{\theta}/\theta-1)^2}{\hat{\theta}/\theta}=4a(\hat{\theta},\theta).
\end{aligned}
$$

(ii) is obvious from (23).□

The inequality (35) clearly implies that $1+\rho(c;m_1,m_2)/4\geq 1+E(a(\hat{\theta},\theta))$. However, the factor $q(r_1,\cdots,r_k;\theta)$ can be so small (or large) that

$$1+\rho(c;m_1,m_2)q(r_1,\cdots,r_k;\theta)\ \leq (or \geq)\ 1+E(a(\hat{\theta},\theta))$$

holds. While the upper bound $1+\rho(c;m_1,m_2)q(r_1,\cdots,r_k;\theta)$ depends on the regressor matrix $\boldsymbol{X}$ through r_i's, the alternative upper bound $1+E(a(\hat{\theta},\theta))$ ignores the information contained in $\boldsymbol{X}$.

Acknowledgement. This research is partially supported by the Ministry of Education in Japan (Kiban (c)(2) No11630028). The author thanks the referee for his/her careful reading and helpful comments.

References

[1] Abramowitz, M. and Stegun, I. (1972), *Handbook of Mathematical Functions*, Dover Publications, New York.

[2] Bilodeau, M. (1990), On the choice of the loss fuction in covariance estimation, *Statist. Decisions.*, **8**, 131-139.

[3] Eaton, M.L. (1985), The Gauss-Markov Theorem in multivariate analysis, in *Multivariate Analysis Y* (P.R.Krishnaiah ed.), 177-201.

[4] Kariya, T. (1981), Bounds for the covariance matrices of Zellner's estimator in the SUR model and 2SAE in a heteroscedasitic model, *Jour. Amer. Statist. Assoc.*, **76**, 975-979.

[5] Kariya, T. (1985), A nonlinear version of the Gauss-Markov Theorem, *Jour. Amer. Statist. Assoc.*, **80**, 476-477.

[6] Kariya, T. and Toyooka, Y. (1985), Nonlinear versions of the Gauss-Markov Theorem and GLSE, in *Multivariate Analysis Y* (P.R. Krishnaiah ed.), 345-354.

[7] Khatri, C.G. and Shah, K.R. (1974), Estimation of location parameters from two linear models under normality, *Comm. Statist. Theory. Methods.*, **3**, 647-663.

[8] Kubokawa, T. (1998), Double shrinkage estimation of common coefficients in two regression equations with heteroscedasticity, *Jour. Multiv. Anal.*, **67**, 169-189.

[9] Kurata, H. and Kariya, T. (1996), Least upper bound for the covariance matrix of a generalized least squares estimator in regression with applications to a seemingly unrelated regression model and a heteroscedastic model, *Ann. Statist.*, **24**, 1547-1559

[10] Kurata, H. (1999), On the efficiencies of several generalized least squares estimators in a seemingly unrelated regression model and a heteroscedastic model, *Jour. Multiv. Anal.*, **70**, 86-94.

[11] Swamy, P.A.V.B. and Mehta, J.S. (1979), Estimation of common coefficients in two regression equations, *Jour. Econometrics.*, **10**, 1-14.

[12] Taylor, W.E. (1977), Small sample properties of a class of two stage Aitken estimators, *Econometrica.*, **45**, 497-508.

[13] Taylor, W.E. (1978), The heteroscedastic linear model: Exact finite sample results, *Econometrica.*, **46**, 663-675.

[14] Toyooka, Y. and Kariya, T. (1986), An approach to upper bound problems for risks of generalized least squares estimators, *Ann. Statist.*, **14**, 679-690.

Postal Address: Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, JAPAN
Telephone Number: 03-5454-6472
E-mail Address: kurata@waka.c.u-tokyo.ac.jp