응용통계연구 제 6권 1호
1993년 pp. 95 ~ 103

Necessary and Sufficient Conditions for the Equality between the Two Best Linear Unbiased Estimators and Their Applications

Sang Ho Lee[3]

Abstract

Necessary and sufficient conditions for the equality between the best linear unbiased estimators in two linear models with different covariance matrices, V_1 and V_2, say, are derived. Various applications of this discovery are also given. Necessary and sufficient conditions for the equality between the best linear unbiased estimator and the ordinary least squares estimator are discussed related to this topic.

1. Introduction

Consider the general linear model

$$Ey = X\beta, \;\; cov(y) = \sigma^2 V \quad , \tag{1}$$

where X is an $n\times p$ matrix of rank p and β is a $p\times 1$ vecter of unknown parameters. The $n\times n$ matrix V is fixed, known, symmetric and positive definite; the constant σ^2 is unknown but plays no role in this article. Therefore, without loss of generality, we may set σ^2=1.

It is known that the best linear unbiased estimator (**BLUE**) $\hat{\beta}_V$ of β in model (1) is unique and turns out to be the generalized least squares estimator (**GLSE**), $\hat{\beta}_V = (X'V^{-1}X)^{-1}X'V^{-1}y$. So when $V = V_1$, the **BLUE** of β can be written as

[3] Department of Statistics, Kangwon National University, Chuncheon, 200-701
Internet : *sanglee@cc.kangwon.ac.kr*

$$\hat{\beta}_{V_1} = (X'V_1^{-1}X)^{-1}X'V_1^{-1}y, \tag{2}$$

and when $V = V_2$, the **BLUE** of β becomes

$$\hat{\beta}_{V_2} = (X'V_2^{-1}X)^{-1}X'V_2^{-1}y\ . \tag{3}$$

We assume that V_1 and V_2 satisfy the above conditions for V. Then what can be the **NASCs** for the equality between $\hat{\beta}_{V_1}$ and $\hat{\beta}_{V_2}$?

2. Derivation of NASCs

We will derive necessary and sufficient conditions (NASCs) for the equality between $\hat{\beta}_{V_1}$ and $\hat{\beta}_{V_2}$. By (2) and (3),

$$\hat{\beta}_{V_1} = \hat{\beta}_{V_2} \text{ iff } (X'V_1^{-1}X)^{-1}X'V_1^{-1}y = (X'V_2^{-1}X)^{-1}X'V_2^{-1}y.$$

Since $X'V_1^{-1}X$ is invertible, this is equivalent to

$$X'V_1^{-1} = X'V_1^{-1}X\ (X'V_2^{-1}X)^{-1}X'V_2^{-1}.$$

In other words,

$$X'V_1^{-1}(I-Q) = 0\ , \tag{4}$$

where $Q = X(X'V_2^{-1}X)^{-1}X'V_2^{-1}$.

Now we claim that (4) is equivalent to $R(V_1^{-1}X) = R(V_2^{-1}X)$, where $R(A)$ denotes the range space of A, i.e., the space generated by the columns of A.

Proof of the above claim

If (4) holds, then the columns of $V_1^{-1}X$ and those of $I-Q$ are orthogonal. Since V_1 is nonsingular, rank$(V_1^{-1}X)$ = rank(X) = p. Since $I-Q$ is idempotent, rank$(I-Q)$ = trace$(I-Q)$ = $n - p$. Therefore, the columns of $V_1^{-1}X$ and those of $I-Q$ together can generate any $n\times 1$ vectors.

It can easily be seen that the columns of $V_2^{-1}X$ and those of $I-Q$ are orthogonal and rank$(V_2^{-1}X)$ = p. So the columns of $V_2^{-1}X$ and those of $I-Q$ also can generate any $n\times 1$ vectors.

Therefore the columns of $V_1^{-1}X$ and those of $V_2^{-1}X$ generate the same space. That is, $\mathbf{R}(V_1^{-1}X) = \mathbf{R}(V_2^{-1}X)$.

Conversely, if $\mathbf{R}(V_1^{-1}X) = \mathbf{R}(V_2^{-1}X)$, then there exists a $p\times p$ nonsingular matrix M such that

$$V_1^{-1}X = V_2^{-1}XM. \tag{5}$$

If we put this relation directly into the BLUE formula, then

$$\begin{aligned}\hat{\beta}_{V_1} &= (X'V_1^{-1}X)^{-1}X'V_1^{-1}y \\ &= (M'X'V_2^{-1}X)^{-1}M'X'V_2^{-1}y \quad \text{by (5)} \\ &= (X'V_2^{-1}X)^{-1}X'V_2^{-1}y \\ &= \hat{\beta}_{V_2}\ .\end{aligned}$$

So the following theorem holds.

Theorem 1 $\hat{\beta}_{V_1} = \hat{\beta}_{V_2}$ if and only if $\mathbf{R}(V_1^{-1}X) = \mathbf{R}(V_2^{-1}X)$ if and only if there exists a $p\times p$ nonsingular matrix M such that $V_1^{-1}X = V_2^{-1}XM$.

3. Remarks

If we let $V_1 = V$ and $V_2 = I$, then this problem is equivalent to finding NASCs for the equality between the BLUE $\hat{\beta}_V$ and the ordinary least squares estimator (OLSE) $\hat{\beta}_I$. In this case, the NASC in theorem 1 becomes $\mathbf{R}(V^{-1}X) = \mathbf{R}(X)$.

In fact, this is a NASC for equality between the BLUE and the OLSE. For its proof, see Kruskal (1968) and Seber (1977). This is equivalent to $\mathbf{R}(VX) = \mathbf{R}(X)$. But in general $\mathbf{R}(V_1X) = \mathbf{R}(V_2X)$ is not a NASC for the equality between the two BLUEs, $\hat{\beta}_{V_1}$ and $\hat{\beta}_{V_2}$.

For other equivalent NASCs for equlity between BLUE and OLSE, see Puntanen and Styan (1989), Zyskind (1967), and many others. Especially Puntanen and Styan (1989) is a good reference for its history. Zyskind (1967), I think, had almost completed it and derived eight NASCs for equality between BLUE and OLSE in the general linear model, where the design matrix X and the covariance matrix V need not be of full column ranks. We, in this article, do not try to find NASCs for equality between BLUEs in the same model as Zyskind had considered. We are rather interested in its application.

4. Application

Several forms of sufficient conditions for the equality between the two BLUEs are used for its application.

Theorem 2

$$\mathbf{R}(V_1 - V_2) \subseteq \mathbf{R}(X) \tag{6}$$

is a sufficient condition but not a necessary condition for $\hat{\beta}_{V_1} = \hat{\beta}_{V_2}$.

Proof(Sufficiency part) If $\mathbf{R}(V_1 - V_2) \subseteq \mathbf{R}(X)$, then there exists a $p \times n$ matrix M_0 such that $V_1 - V_2 = XM_0$. This implies $V_1 = V_2 + XM_0$, and so V_1^{-1} can be written in the form

$$\begin{aligned} V_1^{-1} &= (V_2 + XM_0)^{-1} \\ &= V_2^{-1} - V_2^{-1}X(I + M_0V_2^{-1}X)^{-1}M_0V_2^{-1}. \end{aligned} \tag{7}$$

Thus $V_1^{-1}X = V_2^{-1}X - V_2^{-1}X(I + M_0V_2^{-1}X)^{-1}M_0V_2^{-1}X$. This implies $\mathbf{R}(V_1^{-1}X) \subseteq \mathbf{R}(V_2^{-1}X)$. We can see similarly that $V_2 = V_1 - XM_0$ implies $\mathbf{R}(V_2^{-1}X) \subseteq \mathbf{R}(V_1^{-1}X)$. So, $\mathbf{R}(V_1^{-1}X) = \mathbf{R}(V_2^{-1}X)$, and hence $\hat{\beta}_{V_1} = \hat{\beta}_{V_2}$ by Theorem 1.

(Not necessary part) If we assume that the reverse is true, and if we replace V_2 by $\sigma^2 I$ in (6), then

$$V_1-\sigma^2 I\in \mathbf{R}(X) \tag{8}$$

is a NASC for the equality $\hat{\beta}_{V_1}=\hat{\beta}_I$.

But we can have a counterexample for this. This occurs when $V_1=\sigma^2\mathrm{diag}(1,1,2)$ and $X=(1\ 0\ 0)'$ This satisfies the equality $\hat{\beta}_{V_1}=\hat{\beta}_I$, but not (8). Direct proof for Theorem 2 is also possible using (5).

Corollary 1 $V_1 = V_2 + X\Gamma X'$, for some $p\times p$ symmetric matrix Γ, is a sufficient condition for the equality $\hat{\beta}_{V_1}=\hat{\beta}_{V_2}$.

Proof By Theorem 2, it's enough to prove that

$$\mathbf{R}(V_1-V_2)\subseteq \mathbf{R}(X) \Leftrightarrow V_1=V_2+X\Gamma X',$$

for some $p\times p$ symmetric matrix Γ.

If $\mathbf{R}(V_1-V_2)\subseteq \mathbf{R}(X)$, then there exists a $p\times n$ matrix M_1 such that $V_1-V_2 = XM_1$. Since V_1-V_2 is symmetric, XM_1 is also symmetric. Let $\lambda_1,\cdots,\lambda_n$ be eigenvalues and $v_1,\cdots,v_n$ be their corresponding eigenvectors of XM_1. Without loss of generality, we may set $v_i'v_i=1$ for all $i=1,2,\cdots,n$.

We know that

$$XM_1=\sum_{i=1}^{n}\lambda_i v_i v_i'$$

Let $E_i = v_i v_i'$, then E_i is idempotent and $E_iE_j = 0$ for $i\neq j$.

Therefore,

$$(XM_1)^2=XM_1M_1'X'=\sum\lambda_i^2E_i.$$

If, for some i, $\lambda_i\neq\lambda_j$ for any $j\neq i$, then the E_i is a polynomial in $XM_1M_1'X'$. If there exists $i_1,\cdots,i_k$ such that $\lambda_{i_1}=\cdots=\lambda_{i_k}$, then the $E_{i_1}+\cdots+E_{i_k}$ is a polynomial in $XM_1M_1'X'$.

Therefore $XM_1(=\sum\lambda_i E_i)$ is a polynomial in $XM_1M_1'X'$. That is, $XM_1 = \sum\alpha_k(XM_1M_1'X')^k$. This can be easily represented as $X\Gamma X'$ for some symmetric $p\times p$ matrix Γ.

Conversely, if there exists a symmetric $p\times p$ matrix Γ such that $V_1=V_2+X\Gamma X'$, then it is almost trivial that $\mathbf{R}(V_1-V_2)\subseteq\mathbf{R}(X)$.

Corollary 2 If $V_4=V_3+kUU'$, where k is a constant and U is an $n\times t$ matrix such that $U\in\mathbf{R}(X)$, then $\hat{\beta}_{V_3}=\hat{\beta}_{V_4}$.

Proof kUU' is a special form of $X\Gamma X'$ in Corollary 1.

We will consider a few examples.

Example 1 Consider the simple regression model

$$y_i=\beta_0+\beta_1x_i+\varepsilon_i \quad \text{for } i=1,2,\cdots,n,$$

where

$$cov(\varepsilon_i,\varepsilon_j)=\begin{cases}\sigma_0^2, & i\neq j\\ \sigma_0^2+\sigma_i^2, & i=j.\end{cases}$$

This can be expressed as the model in (1) with

$$y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},\quad X=\begin{pmatrix}1 & x_1\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix},\text{ and } V=V_*+\sigma_0^2 1_n1_n',$$

$$\text{where } V_*=\begin{pmatrix}\sigma_1^2 & & \\ & \vdots & \\ & & \sigma_n^2\end{pmatrix},$$

Since $\sigma_0^2 1_n1_n'$ is a special form of kUU' in Corollary 2, $\hat{\beta}_V=\hat{\beta}_{V_*}$. This means that it makes no difference to introduce σ_0^2 term in the model (1) when we derive the BLUE. For the practical application of this example, see Lee (1991).

Example 2 The following theorem is due to McEloy (1967).

Theorem 3 If trace(V) = $n\sigma^2$ in model (1), then the following two conditions are equivalent:

(i) for any matrix X which has 1_n as its first column,

$$\hat{\beta}_V = \hat{\beta}_I \tag{9}$$

(ii) V can be written in the form

$$V = \sigma^2((1-\rho)I + \rho 1_n 1_n'), \tag{10}$$

where $0 \le \rho \le 1$.

Arbitrariness in (i) of X was not clearly stated in McEloy (1967), so that many authors made mistakes by omitting it when they referred the above theorem in McEloy (1967); see Seber (1977), Odell (1983,p315). Milliken and Albohali (1984) even claimed that McEloy's condition (10) is only sufficient but not necessary. After that, many others assisted McEloy (1967), see Puntanen (1985) and Mathew (1986).

If X is a fixed matrix, as Milliken and Albohali (1984) have thought, then we can easily find a counterexamle. That is, if $V = \sigma^2 I + cc'$, where c is a column in X, which is different from (10), then (9) holds by Corollary 2. So we can easily see that (10) is not necessary for (9) when X is fixed.

Example 3 Consider the following random coefficient model

$$y = X\beta + e\ ,$$

where $\beta \sim N(\mu,\Gamma)$ and $e \sim N(0,V)$. Then this is the same as the model

$$y = X\mu + e\ ,$$

where $e \sim N(0,V_5)$ and $V_5 = V + X\Gamma X'$. The BLUE of μ, $\hat{\beta}_{V_5}$, turns out to be $\hat{\beta}_V = (X'V^{-1}X)^{-1}X'V^{-1}y$ by Corollary 1.

Acknowledgement

The proof in Corollary 1 is due to T. P. Speed who was my thesis advisor during my stay at U.C.-Berkeley.

References

[1] Kruskal, W. (1968), "When are Gauss-Markov and Least Squares Estimators Identical? A Coordinate-Free Approach," *The Annals of Mathematical Statistics,* **39**, 70-75.

[2] Lee, S. H. (1991), *Some Statistical Aspects of Gradient Gel Electrophoresis,* Ph.D. dissertation, University of California at Berkeley, USA.

[3] Mathew, T. (1986), "McEloy's Condition: Comments on Milliken and Albohali and Farebrother," *The American Statistician,* **40**, 179.

[4] McEloy, F. (1967), "A Necessary and Sufficient Condition that Ordinary Least Squares Estimators be Best Linear Unbiased," *Journal of the American Statistical Association,* **62**, 1302-1304.

[5] Milliken, G. A., and Albohali, M. (1984), "On Necessary and Sufficient Conditions for Ordinary Least Squares Estimators to be Best Linear Unbiased Estimators," *The American Statistician,* **38**, 298-299.

[7] Odell, P. L. (1983), "Gauss-Marov Theorem," in *Encyclopedia of Statistical Science* (Vol.3), eds. S. Kotz, N. L. Johnson, and C. B. Read, New York, John Wiley, pp.314-316.

[8] Puntanen, S. (1985), "Further Comment on Milliken and Albohali", *The American Statistician,* **40**, 178.

[9] Puntanen, S., and Styan, G. P. H. (1989), "The Equality of the Ordinary Least Squares Estimator and the Best Linear Unbiased Estimator," *The American Statistician,* **43**, 153-161.

[10] Seber, G. A. F. (1977), *Linear Regression Analysis,* New York: John Wiley.

[11] Zyskind, G.(1967), On Canonical Forms, Non-negative Covariance Matrices and Best and Simple Least Squares Linear Estimators in Linear Models," *The Annals of Mathematical Statistics,* **38**, 1092-1109.

두개의 BLUE가 서로 같을 필요충분조건들과 그 응용

이 상 호[1)]

요 약

두개의 공분산행렬 V_1과 V_2로 구별되는 두개의 선형모형에서 BLUE끼리 같을 필요충분조건이 유도된다. 그리고 이 발견으로 쉽게 이해되는 여러 응용사례도 보여준다. 그동안 여러 논문에서 언급되어 온 BLUE와 OLSE가 같을 필요충분조건도 논의된다.

1) (200-701) 강원도 춘천시 효자2동 강원대학교 통계학과
Internet: sanglee@cc.kangwon.ac.kr