Available online at www.sciencedirect.com
SCIENCE DIRECT®
Statistics & Probability Letters 76 (2006) 1265–1272
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

On equality and proportionality of ordinary least squares, weighted least squares and best linear unbiased estimators in the general linear model

Yongge Tian[a], Douglas P. Wiens[b,*]

[a]School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China
[b]Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alta, Canada T6G 2G1

Received 4 November 2004; received in revised form 23 December 2005
Available online 20 February 2006

Abstract

Equality and proportionality of the ordinary least-squares estimator (OLSE), the weighted least-squares estimator (WLSE), and the best linear unbiased estimator (BLUE) for $\mathbf{X}\boldsymbol{\beta}$ in the general linear (Gauss–Markov) model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{\Sigma}\}$ are investigated through the matrix rank method.

MSC: Primary 62J05; 62H12; 15A24

Keywords: General linear model; Equality; Proportionality; OLSE; WLSE; BLUE; Projectors; Rank formulas for partitioned matrices

1. Introduction

Consider the general linear (Gauss–Markov) model

$$\mathscr{M} = \{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}\}, \tag{1}$$

where $\mathbf{X}$ is a nonnull $n \times p$ known matrix, $\mathbf{y}$ is an $n \times 1$ observable random vector with expectation $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ and with the covariance matrix $Cov(\mathbf{y}) = \boldsymbol{\Sigma}$, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters, and $\boldsymbol{\Sigma}$ is an $n \times n$ symmetric nonnegative definite (n.n.d.) matrix, known entirely except for a positive constant multiplier. In this paper, we investigate some special relations among the ordinary least-squares estimator (OLSE), the weighted least-squares estimator (WLSE), and the best linear unbiased estimator (BLUE) for $\mathbf{X}\boldsymbol{\beta}$ in the model $\mathscr{M}$.

Throughout this paper, $\mathbf{A}'$, $r(\mathbf{A})$ and $\mathscr{R}(\mathbf{A})$ represent the transpose, the rank and the range (column space) of a real matrix $\mathbf{A}$, respectively; $\mathbf{A}^\dagger$ denotes the Moore–Penrose inverse of $\mathbf{A}$ with the properties $\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$, $\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger$, $(\mathbf{A}\mathbf{A}^\dagger)' = \mathbf{A}\mathbf{A}^\dagger$ and $(\mathbf{A}^\dagger\mathbf{A})' = \mathbf{A}^\dagger\mathbf{A}$. In addition, let $\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{A}^\dagger$ and $\mathbf{Q}_\mathbf{A} = \mathbf{I} - \mathbf{P}_\mathbf{A}$.

Let $\mathscr{M}$ be as given in (1). The OLSE $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is defined to be a vector minimizing $\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ and OLSE($\mathbf{X}\boldsymbol{\beta}$) is defined to be $\mathbf{X}\widehat{\boldsymbol{\beta}}$. Suppose $\mathbf{V}$ is an $n \times n$ n.n.d. matrix. The WLSE $\widetilde{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is

*Corresponding author. Tel.: +1 780 4924406; fax: +1 780 4926826.
E-mail addresses: yongge@mail.shufe.edu.cn (Y. Tian), doug.wiens@ualberta.ca (D.P. Wiens).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2006.01.005

defined to be a vector $\boldsymbol{\beta}$ minimizing $\|\mathbf{y}-\mathbf{X}\boldsymbol{\beta}\|_{\mathbf{V}}^2 = (\mathbf{y}-\mathbf{X}\boldsymbol{\beta})'\mathbf{V}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})$, and $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ is defined to be $\mathbf{X}\widetilde{\boldsymbol{\beta}}$. When $\mathbf{V}=\mathbf{I}_n$, $\mathrm{WLSE}_{\mathbf{I}_n}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$. The weight matrix $\mathbf{V}$ in $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ can be constructed from the given matrix $\mathbf{X}$ and $\boldsymbol{\Sigma}$, for example, $\mathbf{V}=\boldsymbol{\Sigma}^{\dagger}$ or $\mathbf{V}=(\mathbf{X}\mathbf{T}\mathbf{X}'+\boldsymbol{\Sigma})^{\dagger}$, see, e.g., Rao (1971). The BLUE of $\mathbf{X}\boldsymbol{\beta}$ is a linear estimator $\mathbf{G}\mathbf{y}$ such that $E(\mathbf{G}\mathbf{y})=\mathbf{X}\boldsymbol{\beta}$ and for any other linear unbiased estimator $\mathbf{M}\mathbf{y}$ of $\mathbf{X}\boldsymbol{\beta}$, $Cov(\mathbf{M}\mathbf{y}) - Cov(\mathbf{G}\mathbf{y}) = \mathbf{M}\boldsymbol{\Sigma}\mathbf{M}' - \mathbf{G}\boldsymbol{\Sigma}\mathbf{G}'$ is nonnegative definite. The three estimators are well known and have been extensively investigated in the literature.

Definition 1. Let $\mathbf{X}$ be an $n \times p$ matrix, and let $\mathbf{V}$ and $\boldsymbol{\Sigma}$ be two $n \times n$ symmetric n.n.d. matrices.

(a) The projector into the range of $\mathbf{X}$ under the seminorm $\|\mathbf{x}\|_{\mathbf{V}}^2 = \mathbf{x}'\mathbf{V}\mathbf{x}$ is defined to be

$$\mathbf{P}_{\mathbf{X}:\mathbf{V}} = \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V} + [\mathbf{X} - \mathbf{X}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X})]\mathbf{U}, \tag{2}$$

where $\mathbf{U}$ is arbitrary.

(b) The BLUE projector $\mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}$ is defined to be the solution $\mathbf{G}$ satisfying the equation

$$\mathbf{G}[\mathbf{X}, \boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}] = [\mathbf{X}, \mathbf{0}],$$

which can be expressed as

$$\mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}} = [\mathbf{X}, \mathbf{0}][\mathbf{X}, \boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}]^{\dagger} + \mathbf{U}_1(\mathbf{I}_n - [\mathbf{X}, \boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}][\mathbf{X}, \boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}]^{\dagger}),$$

where $\mathbf{U}_1$ is arbitrary. The product $\mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$ is unique and can be written as

$$\mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma} = \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}(\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}})^{\dagger}\boldsymbol{\Sigma}. \tag{3}$$

The following results on the OLSE, WLSE and BLUE of $\mathbf{X}\boldsymbol{\beta}$ in $\mathscr{M}$ are well known, see, e.g., Mitra and Rao (1974) and Puntanen and Styan (1989).

Lemma 1. *Let $\mathscr{M}$ be as given in* (1). *Then*:

(a) *The OLSE of $\mathbf{X}\boldsymbol{\beta}$ in $\mathscr{M}$ is unique, and can be written as* $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X}}\mathbf{y}$. *In this case*,

$$E[\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta} \quad \text{and} \quad Cov[\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}}.$$

(b) *The general expression of the WLSE of $\mathbf{X}\boldsymbol{\beta}$ in $\mathscr{M}$ can be written as* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{y}$. *In this case*,

$$E[\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{X}\boldsymbol{\beta} \quad \text{and} \quad Cov[\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_{\mathbf{X}:\mathbf{V}}\boldsymbol{\Sigma}\mathbf{P}'_{\mathbf{X}:\mathbf{V}}.$$

In particular, $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *is unique if and only if* $r(\mathbf{V}\mathbf{X}) = r(\mathbf{X})$. *In such a case*, $\mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{X} = \mathbf{X}$ *and* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *is unbiased.*

(c) *The general expression of the BLUE of $\mathbf{X}\boldsymbol{\beta}$ in $\mathscr{M}$ can be written as* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\mathbf{y}$. *In this case*,

$$E[\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta} \quad \text{and} \quad Cov[\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}\mathbf{P}'_{\mathbf{X}\|\boldsymbol{\Sigma}}.$$

Moreover, $Cov[\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})]$ *can be expressed as*

$$Cov[\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}} - \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}(\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}})^{\dagger}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}}.$$

In particular, $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *is unique if* $\mathbf{y} \in \mathscr{R}[\mathbf{X}, \boldsymbol{\Sigma}]$, *the range of* $[\mathbf{X}, \boldsymbol{\Sigma}]$.

Various properties of $\mathbf{P}_{\mathbf{X}:\mathbf{V}}$ and $\mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}$ can be found in Mitra and Rao (1974), and Puntanen and Styan (1989). Because these estimators are derived from different optimal criteria, they are not necessarily equal. An interesting problem on these estimators is to give necessary and sufficient conditions for them to be equal. If this is true, one can use the OLSE of $\mathbf{X}\boldsymbol{\beta}$ instead of the BLUE of $\mathbf{X}\boldsymbol{\beta}$. Further, it is of interest to consider the proportionality of these estimators.

In order to compare two estimators for an unknown parameter vector, various efficiency functions have been introduced. The most frequently used measure for relations between two unbiased estimators $\mathbf{L}_1\mathbf{y}$ and

$\mathbf{L}_2\mathbf{y}$ of $\boldsymbol{\beta}$ with both $Cov(\mathbf{L}_1\mathbf{y})$ and $Cov(\mathbf{L}_2\mathbf{y})$ nonsingular is the D-relative efficiency

$$\mathrm{eff}_D(\mathbf{L}_1\mathbf{y}, \mathbf{L}_2\mathbf{y}) = \frac{\det[Cov(\mathbf{L}_2\mathbf{y})]}{\det[Cov(\mathbf{L}_2\mathbf{y})]}.$$

A similar relative efficiency function is also defined for the determinants of the information matrices corresponding to two designs for a linear regression model. Other relative efficiency functions can be defined through the traces and norms of covariance matrices, for example,

$$\mathrm{eff}_A(\mathbf{L}_1\mathbf{y}, \mathbf{L}_2\mathbf{y}) = \frac{\mathrm{tr}[Cov(\mathbf{L}_1\mathbf{y})]}{\mathrm{tr}[Cov(\mathbf{L}_2\mathbf{y})]}.$$

If $\mathrm{eff}_\mathrm{D}(\mathbf{L}_1\mathbf{y}, \mathbf{L}_2\mathbf{y}) = 1$, i.e., $\det[Cov(\mathbf{L}_1\mathbf{y})] = \det[Cov(\mathbf{L}_2\mathbf{y})]$, the two estimators $\mathbf{L}_1\mathbf{y}$ and $\mathbf{L}_2\mathbf{y}$ are said to have the same D-efficiency, and is denoted by $\mathbf{L}_1\mathbf{y} \overset{\mathrm{D}}{\sim} \mathbf{L}_2\mathbf{y}$.

In addition to the determinant equality $\det[Cov(\mathbf{L}_1\mathbf{y})] = \det[Cov(\mathbf{L}_2\mathbf{y})]$, some strong relations between two unbiased estimators are defined as follows.

Definition 2. Suppose that $\mathbf{L}_1\mathbf{y}$ and $\mathbf{L}_2\mathbf{y}$ are two unbiased linear estimators of $\boldsymbol{\beta}$ in (1).

(a) The two estimators are said to have the same efficiency if

$$Cov(\mathbf{L}_1\mathbf{y}) = Cov(\mathbf{L}_2\mathbf{y}), \tag{4}$$

and this is denoted by $\mathbf{L}_1\mathbf{y} \overset{\mathrm{C}}{\sim} \mathbf{L}_2\mathbf{y}$.

(b) The two estimators are said to be identical (coincide) with probability 1 if

$$Cov(\mathbf{L}_1\mathbf{y} - \mathbf{L}_2\mathbf{y}) = \mathbf{0}, \tag{5}$$

and this is denoted by $\mathbf{L}_1\mathbf{y} \overset{\mathrm{P}}{\sim} \mathbf{L}_2\mathbf{y}$.

Clearly (4) is equivalent to

$$\mathbf{L}_1\boldsymbol{\Sigma}\mathbf{L}_1' = \mathbf{L}_2\boldsymbol{\Sigma}\mathbf{L}_2', \tag{6}$$

and (5) is equivalent to $(\mathbf{L}_1 - \mathbf{L}_2)\boldsymbol{\Sigma}(\mathbf{L}_1 - \mathbf{L}_2)' = \mathbf{0}$, i.e.,

$$\mathbf{L}_1\boldsymbol{\Sigma} = \mathbf{L}_2\boldsymbol{\Sigma}. \tag{7}$$

It is easy to see from (6) and (7) that

$$\mathbf{L}_1\mathbf{y} \overset{\mathrm{P}}{\sim} \mathbf{L}_2\mathbf{y} \Rightarrow \mathbf{L}_1\mathbf{y} \overset{\mathrm{C}}{\sim} \mathbf{L}_2\mathbf{y} \Rightarrow \mathbf{L}_1\mathbf{y} \overset{\mathrm{D}}{\sim} \mathbf{L}_2\mathbf{y}.$$

However, the reverse implication is not true. For example, let

$$\boldsymbol{\Sigma} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{L}_1 = [1, -1], \quad \mathbf{L}_2 = [-1, 1].$$

Then

$$\mathbf{L}_1\boldsymbol{\Sigma}\mathbf{L}_1' = \mathbf{L}_2\boldsymbol{\Sigma}\mathbf{L}_2' = 2, \quad \text{but } (\mathbf{L}_1 - \mathbf{L}_2)\boldsymbol{\Sigma} = [2, -2] \neq \mathbf{0}.$$

Many authors investigated the equalities among these estimators using Definition 2(a); see, e.g., Baksalary and Puntanen (1989), Puntanen and Styan (1989), Young et al. (2000), and Groß et al. (2001) among others. In this paper, we use Definition 2(b) to characterize the relations between two estimators.

Lemma 2. *Let* $\mathscr{M}$ *be as given in* (1), *and suppose* $\mathbf{L}_1\mathbf{y}$ *and* $\mathbf{L}_2\mathbf{y}$ *are two linear estimators for* $\mathbf{X}\boldsymbol{\beta}$ *in* $\mathscr{M}$. *Then* $\mathbf{L}_1\mathbf{y} = \mathbf{L}_2\mathbf{y}$ *with probability* 1, *i.e.,* $\mathbf{L}_1\mathbf{y} \overset{\mathrm{P}}{\sim} \mathbf{L}_2\mathbf{y}$, *if and only if* $\mathbf{L}_1\mathbf{X} = \mathbf{L}_2\mathbf{X}$ *and* $\mathbf{L}_1\boldsymbol{\Sigma} = \mathbf{L}_2\boldsymbol{\Sigma}$, *i.e.,* $\mathbf{L}_1[\mathbf{X}, \boldsymbol{\Sigma}] = \mathbf{L}_2[\mathbf{X}, \boldsymbol{\Sigma}]$.

The rank of a matrix is defined to be the dimension of the column space or row space of the matrix. It has been recognized since the 1970s that rank equalities for matrices provide a powerful method for finding general properties of matrix expressions. In fact, for any two matrices A and B of the same size, the equality

$A = B$ holds if and only if $r(A - B) = 0$; two sets S_1 and S_2 consisting of matrices of the same size have a common matrix if and only if $\min_{A \in S_1, B \in S_2} r(A - B) = 0$; the set inclusion $S_1 \subseteq S_2$ holds if and only if $\max_{A \in S_1} \min_{B \in S_2} r(A - B) = 0$. If some formulas for the rank of $A - B$ are derived, they can be used to characterize the equality $A = B$. In order to simplify various matrix expressions involving generalized inverses, we need some rank equalities for partitioned matrices due to Marsaglia and Styan (1974).

Lemma 3. *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$ *and* $C \in \mathbb{R}^{l \times n}$. *Then*

$$r[\mathbf{A}, \mathbf{B}] = r(\mathbf{A}) + r[(\mathbf{I}_m - \mathbf{A}\mathbf{A}^-)\mathbf{B}] = r(\mathbf{B}) + r[(\mathbf{I}_m - \mathbf{B}\mathbf{B}^-)\mathbf{A}], \tag{8}$$

$$r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) + r[\mathbf{C}(\mathbf{I}_n - \mathbf{A}^-\mathbf{A})] = r(\mathbf{C}) + r[\mathbf{A}(\mathbf{I}_n - \mathbf{C}^-\mathbf{C})], \tag{9}$$

$$r\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix} = r(\mathbf{B}) + r(\mathbf{C}) + r[(\mathbf{I}_m - \mathbf{B}\mathbf{B}^-)\mathbf{A}(\mathbf{I}_n - \mathbf{C}^-\mathbf{C})]. \tag{10}$$

The following result is shown in Tian (2002), and Tian and Cheng (2003).

Lemma 4. *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$ *and* $\mathbf{C} \in \mathbb{R}^{l \times n}$. *Then*

$$\min_{\mathbf{X} \in \mathbb{R}^{k \times l}} r(\mathbf{A} - \mathbf{B}\mathbf{X}\mathbf{C}) = r[\mathbf{A}, \mathbf{B}] + r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} - r\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0} \end{bmatrix}. \tag{11}$$

2. Equality and proportionality of estimators

Theorem 1. *Let* $\mathscr{M}$ *be as given in* (1).

(a) *There is* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *such that* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *if and only if* $\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{0}$.

(b) *Assume* $r[\mathbf{A}, \boldsymbol{\Sigma}] = n$. *Then there is* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *such that* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *if and only if* $\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{V}\mathbf{P}_{\mathbf{X}}$.

Proof. From Lemma 2, we see that $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ if and only if there is $\mathbf{P}_{\mathbf{X}:\mathbf{V}}$ such that

$$\mathbf{P}_{\mathbf{X}:\mathbf{V}}[\mathbf{X}, \boldsymbol{\Sigma}] = [\mathbf{X}, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}].$$

Let $\mathbf{M}_1 = [\mathbf{X}, \boldsymbol{\Sigma}]$ and $\mathbf{M}_2 = [\mathbf{X}, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}]$. Then by (2)

$$\mathbf{P}_{\mathbf{X}:\mathbf{V}}[\mathbf{X}, \boldsymbol{\Sigma}] - [\mathbf{X}, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}] = \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 - (\mathbf{X} - \mathbf{A}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X}))\mathbf{U}\mathbf{M}_1,$$

where $\mathbf{U}$ is arbitrary. From (11),

$$\begin{aligned} \min_{\mathbf{P}_{\mathbf{X}:\mathbf{V}}} r(\mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{M}_1 - \mathbf{M}_2) = {} & \min_{\mathbf{U}} r[\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 - (\mathbf{X} - \mathbf{A}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X}))\mathbf{U}\mathbf{M}_1] \\ & = r[\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2, \mathbf{X} - \mathbf{X}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X})] \\ & \quad + r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 \\ \mathbf{M}_1 \end{bmatrix} \\ & \quad - r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 & \mathbf{X} - \mathbf{X}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X}) \\ \mathbf{M}_1 & \mathbf{0} \end{bmatrix}. \end{aligned} \tag{12}$$

By (8)–(10) and elementary block matrix operations,

$$
\begin{aligned}
r[\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2, \mathbf{X} - \mathbf{X}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X})] &= r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 & \mathbf{X} \\ \mathbf{0} & \mathbf{V}\mathbf{X} \end{bmatrix} - r(\mathbf{V}\mathbf{X}) \\
&= r\begin{bmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{V}\mathbf{M}_2 & \mathbf{0} \end{bmatrix} - r(\mathbf{V}\mathbf{X}) \\
&= r[\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{V}\mathbf{M}_2] + r(\mathbf{X}) - r(\mathbf{V}\mathbf{X}),
\end{aligned}
$$

$$
r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 \\ \mathbf{M}_1 \end{bmatrix} = r\begin{bmatrix} \mathbf{M}_2 \\ \mathbf{M}_1 \end{bmatrix} = r\begin{bmatrix} \mathbf{X} & \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} \\ \mathbf{X} & \boldsymbol{\Sigma} \end{bmatrix} = r\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{X} & \boldsymbol{\Sigma} \end{bmatrix} = r(\mathbf{M}_1),
$$

$$
\begin{aligned}
r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{M}_2 & \mathbf{X} - \mathbf{X}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X}) \\ \mathbf{M}_1 & \mathbf{0} \end{bmatrix} &= r(\mathbf{M}_1) + r[\mathbf{X} - \mathbf{X}(\mathbf{V}\mathbf{X})^{\dagger}(\mathbf{V}\mathbf{X})] \\
&= r(\mathbf{M}_1) + r(\mathbf{X}) - r(\mathbf{V}\mathbf{X}).
\end{aligned}
$$

Also note that $\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{X} = \mathbf{V}\mathbf{X}$. Hence,

$$\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{V}\mathbf{M}_2 = [\mathbf{0}, \mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\boldsymbol{\Sigma} - \mathbf{V}\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}].$$

Moreover,

$$\mathbf{P}_{\mathbf{X}}[\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\boldsymbol{\Sigma} - \mathbf{V}\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}] = \mathbf{P}_{\mathbf{X}}\mathbf{V}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}.$$

Hence, $r[\mathbf{V}\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \mathbf{V}\mathbf{M}_2] = r(\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma})$. Substituting these rank equalities into (12) and simplifying gives

$$\min_{\mathbf{P}_{\mathbf{X}:\mathbf{V}}} r(\mathbf{P}_{\mathbf{X}:\mathbf{V}}[\mathbf{X}, \boldsymbol{\Sigma}] - [\mathbf{X}, \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}]) = r(\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}).$$

Let the right-hand side be zero, we obtain the result in (a). The rank equality $r[\mathbf{X}, \boldsymbol{\Sigma}] = n$ is equivalent to $r(\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}) = r(\mathbf{Q}_{\mathbf{X}})$ by (8). Hence $\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{0}$ is equivalent to $\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}} = \mathbf{0}$, i.e., $\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{P}_{\mathbf{X}}$. Note that $\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{P}_{\mathbf{X}}$ is symmetric, $\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{P}_{\mathbf{X}}$ is equivalent to $\mathbf{P}_{\mathbf{X}}\mathbf{V} = \mathbf{V}\mathbf{P}_{\mathbf{X}}$. □

Some other problems on $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ can be investigated. For example, one can solve for n.n.d. $\mathbf{V}$ from the equation $\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{0}$ if both $\mathbf{X}$ and $\boldsymbol{\Sigma}$ are given, or solve for n.n.d. $\boldsymbol{\Sigma}$ from the equation if both $\mathbf{X}$ and $\mathbf{V}$ are given.

Theorem 2. *Let* $\mathscr{M}$ *be as given in* (1). *Then the following statements are equivalent*:

(a) *There is* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *such that* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$.
(b) *There is* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *such that* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{C}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$.
(c) $\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}} = \mathbf{0}$, *i.e.*, $\mathscr{R}(\boldsymbol{\Sigma}\mathbf{X}) \subseteq \mathscr{R}(\mathbf{X})$.
(d) $\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} = \boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}}$.

Proof. From Definition 1(a) and (c) and Lemma 2, we see that there is $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ such that $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ if and only if $\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} = \mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma}$. From (3), we obtain

$$\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma} - \mathbf{P}_{\mathbf{X}\|\boldsymbol{\Sigma}}\boldsymbol{\Sigma} = \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}(\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}})^{\dagger}\boldsymbol{\Sigma}.$$

It is easy to verify that

$$r[\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}(\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}})^{\dagger}\boldsymbol{\Sigma}] = r(\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}) = r(\mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}}).$$

The equivalence of (a), (c) and (d) follows from this result. From Lemma 1(a) and (c), we see that

$$Cov[\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})] - Cov[\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{P}_{\mathbf{X}}\boldsymbol{\Sigma}(\mathbf{Q}_{\mathbf{X}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathbf{X}})^{\dagger}\boldsymbol{\Sigma}\mathbf{P}_{\mathbf{X}}.$$

It is easy to verify that

$$r[\mathbf{P_X\Sigma(Q_X\Sigma Q_X)^\dagger\Sigma P_X}] = r[\mathbf{P_X\Sigma(Q_X\Sigma Q_X)^\dagger}] = r(\mathbf{P_X\Sigma Q_X\Sigma Q_X}) = r(\mathbf{P_X\Sigma Q_X}).$$

The equivalence of (b) and (d) follows from this result. □

Theorem 2(b)–(d) are well known, see, e.g., Puntanen and Styan (1989), Young et al. (2000).

Theorem 3. *Let $\mathscr{M}$ be as given in* (1).

(a) *There are* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *and* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *such that* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *if and only if* $\mathbf{P_X V\Sigma Q_X} = \mathbf{0}$, *that is*, $\mathscr{R}(\mathbf{\Sigma V X}) \subseteq \mathscr{R}(\mathbf{X})$.
(b) *If* $\mathbf{V} = (\mathbf{XX}' + \mathbf{\Sigma})^\dagger$, *then* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *is unique and* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$.

Proof. From Lemma 2, $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ if and only if

$$\mathbf{P}_{\mathbf{X}:\mathbf{V}}[\mathbf{X}, \mathbf{\Sigma}] = [\mathbf{X}, \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}].$$

Let $\mathbf{M}_1 = [\mathbf{X}, \mathbf{\Sigma}]$ and $\mathbf{M}_2 = [\mathbf{X}, \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}]$. Then by (11),

$$\begin{aligned}\min_{\mathbf{P}_{\mathbf{X}:\mathbf{V}}} r(\mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{M}_1 - \mathbf{M}_2) &= \min_{\mathbf{U}} r[\mathbf{X(X'VX)^\dagger X'VM_1 - M_2 - (X - X(VX)^\dagger(VX))UM_1}] \\ &= r[\mathbf{X(X'VX)^\dagger X'VM_1 - M_2, X - X(VX)^\dagger(VX)}] \\ &\quad + r\begin{bmatrix} \mathbf{X(X'VX)^\dagger X'VM_1 - M_2} \\ \mathbf{M_1} \end{bmatrix} \\ &\quad - r\begin{bmatrix} \mathbf{X(X'VX)^\dagger X'VM_1 - M_2} & \mathbf{X - X(VX)^\dagger(VX)} \\ \mathbf{M_1} & \mathbf{0} \end{bmatrix}. \end{aligned} \tag{13}$$

By (8)–(10),

$$\begin{aligned} r[\mathbf{X(X'VX)^\dagger X'VM_1 - M_2, X - X(VX)^\dagger(VX)}] &= r\begin{bmatrix} \mathbf{X(X'VX)^\dagger X'VM_1 - M_2} & \mathbf{X} \\ \mathbf{0} & \mathbf{VX} \end{bmatrix} - r(\mathbf{VA}) \\ &= r\begin{bmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{VV(X'VX)^\dagger X'VM_1 - VM_2} & \mathbf{0} \end{bmatrix} - r(\mathbf{VX}) \\ &= r[\mathbf{VX(X'VX)^\dagger X'VM_1 - VM_2}] + r(\mathbf{X}) - r(\mathbf{VX}), \end{aligned}$$

$$r\begin{bmatrix} \mathbf{X(X'VX)^\dagger X'VM_1 - M_2} \\ \mathbf{M_1} \end{bmatrix} = r\begin{bmatrix} \mathbf{M_2} \\ \mathbf{M_1} \end{bmatrix} = r\begin{bmatrix} \mathbf{0} \\ \mathbf{M_1} \end{bmatrix} = r(\mathbf{M}_1),$$

$$\begin{aligned} r\begin{bmatrix} \mathbf{X(X'VX)^\dagger X'VM_1 - M_2} & \mathbf{X - X(VX)^\dagger(VX)} \\ \mathbf{M_1} & \mathbf{0} \end{bmatrix} &= r(\mathbf{M}_1) + r[\mathbf{X - X(VX)^\dagger(VX)}] \\ &= r(\mathbf{M}_1) + r(\mathbf{X}) - r(\mathbf{VX}). \end{aligned}$$

Also note that $\mathbf{VX(X'VX)^\dagger X'VX = VX}$. Hence,

$$\mathbf{VX(X'VX)^\dagger X'VM_1 - VM_2} = [\mathbf{0}, \mathbf{VX(X'VX)^\dagger X'V\Sigma} - \mathbf{VP}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}].$$

It can be verified that

$$r[\mathbf{VX(X'VX)^\dagger X'V\Sigma} - \mathbf{VP}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}] = r(\mathbf{P_X V\Sigma Q_X}).$$

Substituting these rank equalities into (13) and simplifying leads to

$$\min_{\mathbf{P}_{\mathbf{X}:\mathbf{V}}} r(\mathbf{P}_{\mathbf{X}:\mathbf{V}}[\mathbf{X}, \mathbf{\Sigma}] - [\mathbf{X}, \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}]) = r(\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{\Sigma}\mathbf{Q}_{\mathbf{X}}),$$

completing the proof. □

The following results are concerned with the proportionality of the three estimators.

Theorem 4. *Let $\mathscr{M}$ be as given in* (1) *and suppose that* $\mathbf{VX}\neq\mathbf{0}$. *Then there are* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *and a scalar* λ *such that* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \lambda\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *if and only if* $\lambda = 1$ *and* $\mathbf{P}_{\mathbf{X}}\mathbf{V}\mathbf{Q}_{\mathbf{X}}\mathbf{\Sigma} = \mathbf{0}$, *that is, there is* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *proportional to* $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *with probability* 1 *if and only if the* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *satisfies* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$.

Proof. From Lemma 2, there is $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ so that $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \lambda\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ if and only if

$$\mathbf{P}_{\mathbf{X}:\mathbf{V}}[\mathbf{X}, \mathbf{\Sigma}] = [\lambda\mathbf{X}, \lambda\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma}].$$

Let $\mathbf{M}_1 = [\mathbf{X}, \mathbf{\Sigma}]$ and $\mathbf{M}_2 = [\mathbf{X}, \mathbf{P}_{\mathbf{X}}\mathbf{\Sigma}]$. Then we can find through the matrix rank method that

$$\min_{\mathbf{P}_{\mathbf{X}:\mathbf{V}}} r(\mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{M}_1 - \lambda\mathbf{M}_2) = r[(1-\lambda)\mathbf{VX}, \mathbf{VX}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{\Sigma} - \lambda\mathbf{V}\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma}]. \tag{14}$$

Let the right-hand side of (14) be zero, we see that $\lambda = 1$ and $\mathbf{VX}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{\Sigma} = \mathbf{V}\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma}$. Hence, the result in this theorem follows from Theorem 1. □

Theorem 5. *Let $\mathscr{M}$ be as given in* (1) *and suppose that* $\mathbf{X}\neq\mathbf{0}$. *Then there are* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *and scalar* λ *such that* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \lambda\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *if and only if* $\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma} = \mathbf{\Sigma}\mathbf{P}_{\mathbf{X}}$, *that is, there is* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *proportional to* $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ *with probability* 1 *if and only if the* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *satisfies* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$.

Proof. From Lemma 2, $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \lambda\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta})$ holds, if and only if

$$\mathbf{X} = \lambda\mathbf{X} \quad \text{and} \quad \mathbf{\Sigma}\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma} = \lambda\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma},$$

which are equivalent to $\lambda=1$ and $\mathbf{\Sigma}\mathbf{P}_{\mathbf{X}}\mathbf{\Sigma} = \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}$. Hence, the result in this theorem follows from Theorem 2. □

Theorem 6. *Let $\mathscr{M}$ be as given in* (1) *and suppose that* $\mathbf{VX}\neq\mathbf{0}$. *Then there are* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *and* λ *such that* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \lambda\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *if and only if* $\lambda = 1$ *and* $\mathscr{R}(\mathbf{\Sigma}\mathbf{VX}) \subseteq \mathscr{R}(\mathbf{X})$, *that is, there is* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *proportional to* $\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ *with probability* 1 *if and only if the* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta})$ *satisfies* $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$.

Proof. From Lemma 2, $\mathrm{WLSE}_{\mathbf{V}}(\mathbf{X}\boldsymbol{\beta}) \overset{\mathrm{P}}{\sim} \lambda\mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ if and only if

$$\mathbf{P}_{A\mathbf{X}:\mathbf{V}}[\mathbf{X}, \mathbf{\Sigma}] = \lambda[\mathbf{X}, \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}].$$

Let $\mathbf{M}_1 = [\mathbf{X}, \mathbf{\Sigma}]$ and $\mathbf{M}_2 = [\mathbf{X}, \mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}]$. Then by (11),

$$\begin{aligned}
\min_{\mathbf{P}_{\mathbf{X}:\mathbf{V}}} r(\mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{M}_1 - \lambda\mathbf{M}_2) &= \min_{\mathbf{U}} r[\mathbf{X}(\mathbf{X}'\mathbf{VA})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \lambda\mathbf{M}_2 - (\mathbf{X} - \mathbf{X}(\mathbf{VA})^{\dagger}(\mathbf{VA}))\mathbf{U}\mathbf{M}_1] \\
&= r[\mathbf{X}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \lambda\mathbf{M}_2, \mathbf{X} - \mathbf{X}(\mathbf{VX})^{\dagger}(\mathbf{VX})] \\
&\quad + r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \lambda\mathbf{M}_2 \\ \mathbf{M}_1 \end{bmatrix} \\
&\quad - r\begin{bmatrix} \mathbf{X}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \lambda\mathbf{M}_2 & \mathbf{X} - \mathbf{X}(\mathbf{VX})^{\dagger}(\mathbf{VX}) \\ \mathbf{M}_1 & \mathbf{0} \end{bmatrix} \\
&= r[\mathbf{VX}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{M}_1 - \lambda\mathbf{V}\mathbf{M}_2] \\
&= r[(1-\lambda)\mathbf{VX}, \mathbf{VX}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{\Sigma} - \lambda\mathbf{V}\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}].
\end{aligned} \tag{15}$$

Let the right-hand side of (15) be zero, we obtain that $\lambda = 1$ and $\mathbf{VX}(\mathbf{X}'\mathbf{VX})^{\dagger}\mathbf{X}'\mathbf{V}\mathbf{\Sigma} = \mathbf{V}\mathbf{P}_{\mathbf{X}\|\mathbf{\Sigma}}\mathbf{\Sigma}$. Hence, the result in this theorem follows from Theorem 3. □

It can be seen from Theorems 4–6 that if any two of the OLSE, WLSE and BLUE for $\mathbf{X}\boldsymbol{\beta}$ in the model (1) are proportional with probability 1, the two estimators are identical with probability 1.

Many other problems on the model (1) can be investigated through the matrix rank method. For example, assume that the general linear model $\{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma}\}$ incorrectly specifies the covariance matrix as $\sigma^2\boldsymbol{\Sigma}_0$, where $\boldsymbol{\Sigma}_0$ is a given n.n.d. matrix and σ^2 is a positive parameter (possibly unknown). Then consider the relations between the BLUEs of $\mathbf{X}\boldsymbol{\beta}$ in the original model and the misspecified model $\{\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{\Sigma}_0\}$. Some previous results on the relations between the BLUEs of $\mathbf{X}\boldsymbol{\beta}$ in the two models can be found in Mitra and Moore (1973), and Mathew (1983). In addition, it is of interest to consider partial coincidence and partial proportionality of the OLSE, WLSE and BLUE of $\mathbf{X}\boldsymbol{\beta}$ in (1), as well as to consider coincidence and proportionality of the OLSE, WLSE and BLUE of $\mathbf{X}\boldsymbol{\beta}$ in (1) under a restriction $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$.

Acknowledgment

The research of both authors was supported by the Natural Sciences and Engineering Research Council of Canada.

References

Baksalary, J.K., Puntanen, S., 1989. Weighted-least squares estimation in the general Gauss–Markov model. In: Dodge, Y. (Ed.), Statistical Data Analysis and Inference. Elsevier, New York, pp. 355–368.

Groß, G., Trenkler, T., Werner, H.J., 2001. The equality of linear transformations of the ordinary least squares estimator and the best linear unbiased estimator. Sankhyā Ser. A 63, 118–127.

Marsaglia, G., Styan, G.P.H., 1974. Equalities and inequalities for ranks of matrices. Linear Multilinear Algebra 2, 269–292.

Mathew, T., 1983. Linear estimation with an incorrect dispersion matrix in linear models with a common linear part. J. Amer. Statist. Assoc. 78, 468–471.

Mitra, S.K., Moore, B.J., 1973. Gauss-Markov estimation with an incorrect dispersion matrix. Sankhyā Ser. A 35, 139–152.

Mitra, S.K., Rao, C.R., 1974. Projections under seminorms and generalized Moore–Penrose inverses. Linear Algebra Appl. 9, 155–167.

Puntanen, S., Styan, G.P.H., 1989. The equality of the ordinary least squares estimator and the best linear unbiased estimator. With comments by O. Kempthorne, S.R. Searle, and a reply by the authors. Amer. Statist. 43, 153–164.

Rao, C.R., 1971. Unified theory of linear estimation. Sankhyā Ser. A 33, 371–394.

Tian, Y., 2002. The maximal and minimal ranks of some expressions of generalized inverses of matrices. Southeast Asian Bull. Math. 25, 745–755.

Tian, Y., Cheng, S., 2003. The maximal and minimal ranks of $A - BXC$ with applications. New York J. Math. 9, 345–362.

Young, D.M., Odell, P.L., Hahn, W., 2000. Nonnegative-definite covariance structures for which the blu, wls, and ls estimators are equal. Statist. Probab. Lett. 49, 271–276.