Communications in Statistics—Theory and Methods, 37: 55–69, 2008

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920701648862

Taylor & Francis
Taylor & Francis Group

Linear Models and Regression Analysis

On Sum Decompositions of Weighted Least-Squares Estimators for the Partitioned Linear Model

YONGGE TIAN[1] AND YOSHIO TAKANE[2]

[1]School of Economics, Shanghai University of Finance and Economics, Shanghai, China
[2]Department of Psychology, McGill University, Montréal, Québec, Canada

For the partitioned linear model $\mathscr{M} = \{\mathbf{y}, \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2, \sigma^2\Sigma\}$, this article investigates decompositions of weighted least-squares estimator (WLSE) of $\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2$ under $\mathscr{M}$ as sums of WLSEs under the two small models $\{\mathbf{y}, \mathbf{X}_1\boldsymbol{\beta}_1, \sigma^2\Sigma\}$ and $\{\mathbf{y}, \mathbf{X}_2\boldsymbol{\beta}_2, \sigma^2\Sigma\}$. Some consequences on the sum decomposition of the unique best unbiased linear estimator (BLUE) of $\mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2$ under $\mathscr{M}$ are also given.

Keywords Matrix rank method; Moore–Penrose inverse; Partitioned linear model; Projector; Small models; Sum decompositions; WLSE.

Mathematics Subject Classification Primary 62J05; Secondary 62H12.

1. Introduction

Throughout this article, $\mathbb{R}^{m\times n}$ stands for the collection of all $m \times n$ real matrices, and the symbols $\mathbf{A}'$, $r(\mathbf{A})$, and $\mathscr{R}(\mathbf{A})$ stand for the transpose, rank, and range (column space) of a matrix $\mathbf{A}$, respectively. A pair of matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{m\times p}$ are said to be orthogonal with respect to a non-negative definite matrix $\mathbf{V} \in \mathbb{R}^{m\times m}$, or $\mathbf{V}$-orthogonal, if $\mathbf{A}'\mathbf{V}\mathbf{B} = \mathbf{0}$. For a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, the Moore–Penrose inverse of $\mathbf{A}$, denoted by $\mathbf{A}^+$, is defined to be the unique solution $\mathbf{X} \in \mathbb{R}^{n\times m}$ to the four Penrose equations:

$$\text{(i) } \mathbf{AXA} = \mathbf{A}, \quad \text{(ii) } \mathbf{XAX} = \mathbf{X}, \quad \text{(iii) } (\mathbf{AX})' = \mathbf{AX}, \quad \text{(iv) } (\mathbf{XA})' = \mathbf{XA}.$$

A matrix $\mathbf{X} \in \mathbb{R}^{n\times m}$ is called a g-inverse of $\mathbf{A}$, denoted by $\mathbf{A}^-$, if it satisfies (i); called an outer inverse of $\mathbf{A}$ if it satisfies (ii). Further, let $\mathbf{P}_\mathbf{A}$, $\mathbf{E}_\mathbf{A}$, and $\mathbf{F}_\mathbf{A}$ stand for the

Received June 30, 2006; Accepted February 28, 2007
Address correspondence to Yongge Tian, School of Economics, Shanghai University of Finance and Economics, Shanghai 200433, China; E-mail: yongge@mail.shufe.edu.cn

three orthogonal projectors $\mathbf{P_A} = \mathbf{AA}^+$, $\mathbf{E_A} = \mathbf{I} - \mathbf{P_A} = \mathbf{I} - \mathbf{AA}^+$, and $\mathbf{F_A} = \mathbf{I} - \mathbf{P_{A'}} = \mathbf{I} - \mathbf{A}^+\mathbf{A}$.

Suppose we are given a partitioned linear model

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \varepsilon, \quad E(\varepsilon) = \mathbf{0}, \quad Cov(\varepsilon) = \sigma^2\Sigma, \tag{1.1}$$

where $\mathbf{X}_1 \in \mathbb{R}^{n\times p_1}$ and $\mathbf{X}_2 \in \mathbb{R}^{n\times p_2}$ are two known matrices of arbitrary rank with $p_1 + p_2 = p$, $\mathbf{y} \in \mathbb{R}^{n\times 1}$ is an observable random vector, $\boldsymbol{\beta}_1 \in \mathbb{R}^{p_1\times 1}$ and $\boldsymbol{\beta}_2 \in \mathbb{R}^{p_2\times 1}$ are two vectors of unknown parameters to be estimated, $\Sigma \in \mathbb{R}^{n\times n}$ is a known nonnegative definite matrix of arbitrary rank, and σ^2 is an unknown positive parameter. If Σ is a singular matrix, (1.1) is also said to be a singular linear model.

The model in (1.1) is often written as a triplet:

$$\mathscr{M} = \{\mathbf{y},\ \mathbf{X}\boldsymbol{\beta},\ \sigma^2\Sigma\} = \{\mathbf{y},\ \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2,\ \sigma^2\Sigma\}, \tag{1.2}$$

where $\mathbf{X} = [\mathbf{X}_1,\ \mathbf{X}_2]$ and $\boldsymbol{\beta} = [\boldsymbol{\beta}_1',\ \boldsymbol{\beta}_2']'$. Of particular interest in the investigation of a partitioned model is relationships between the partitioned model (full model) and its various small or reduced models. This subject was widely investigated from various aspects, see, e.g., Bhimasankaram and Saharay (1997), Chu et al. (2004), Groß and Puntanen (2000), Nurhonen and Puntanen (1992), Werner and Yapar (1995, 1996), and Zhang et al. (2004). For the full model in (1.2), the two small linear models are given by

$$\mathscr{M}_1 = \{\mathbf{y},\ \mathbf{X}_1\boldsymbol{\beta}_1,\ \sigma^2\Sigma\} \quad \text{and} \quad \mathscr{M}_2 = \{\mathbf{y},\ \mathbf{X}_2\boldsymbol{\beta}_2,\ \sigma^2\Sigma\}. \tag{1.3}$$

It is well known that if $\mathbf{X}$ has full column rank, then the ordinary least-squares estimator (OLSE) of $\mathbf{X}\boldsymbol{\beta}$ under (1.2) can be written as $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. If the two submatrices $\mathbf{X}_1$ and $\mathbf{X}_2$ in $\mathbf{X}$ are orthogonal, that is, $\mathbf{X}_1'\mathbf{X}_2 = \mathbf{0}$, then $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ can directly be written as the sum

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' = \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1' + \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'. \tag{1.4}$$

Correspondingly, the $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ can be decomposed as

$$\begin{aligned}\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) &= \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y} + \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{y}\\ &= \mathrm{OLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{OLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2).\end{aligned} \tag{1.5}$$

This equality implies that the OLSE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$ can be written the sum of the two OLSEs under the two small models in (1.3) if $\mathbf{X}_1'\mathbf{X}_2 = \mathbf{0}$. This simple property prompts us to consider decompositions of some other estimators under $\mathscr{M}$ as sums of estimators under the two small models in (1.3). The main purpose of the present article is to extend the equality in (1.5) to weighted least-squares estimators of $\mathbf{X}\boldsymbol{\beta}$ under (1.2).

Let $\mathbf{V} \in \mathbb{R}^{n\times n}$ be a non-negative definite matrix, i.e., $\mathbf{V}$ can be written as $\mathbf{V} = \mathbf{Z}\mathbf{Z}'$ for some matrix $\mathbf{Z}$. The seminorm of a vector $\mathbf{x} \in \mathbb{R}^{n\times 1}$ induced by the weight matrix $\mathbf{V}$ is defined by $\|\mathbf{x}\|_{\mathbf{V}} = (\mathbf{x}'\mathbf{V}\mathbf{x})^{1/2}$. The weighted least-squares estimator (WLSE) of $\boldsymbol{\beta}$ under $\mathscr{M}$ in (1.2) is defined to be

$$\tilde{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}\ \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_{\mathbf{V}}^2. \tag{1.6}$$

The normal matrix equation corresponding to (1.6) is $\mathbf{X}'\mathbf{V}\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{V}\mathbf{y}$. This equation is always consistent. Solving this equation gives the following well-known result.

Lemma 1.1. *The general expression of the WLSE of $\boldsymbol{\beta}$ under $\mathscr{M}$ is*

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}\mathbf{y} + [\mathbf{I}_p - (\mathbf{V}\mathbf{X})^{+}(\mathbf{V}\mathbf{X})]\mathbf{v} = (\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}\mathbf{y} + \mathbf{F}_{\mathbf{V}\mathbf{X}}\mathbf{u}, \tag{1.7}$$

where $\mathbf{u} \in \mathbb{R}^{p\times 1}$ is arbitrary.

For $\mathbf{y} \neq \mathbf{0}$, let $\mathbf{u} = \mathbf{U}\mathbf{y}$ in (1.7), where $\mathbf{U} \in \mathbb{R}^{p\times n}$ is arbitrary. Then (1.7) can be rewritten as the following homogeneous form $\tilde{\boldsymbol{\beta}} = [(\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V} + \mathbf{F}_{\mathbf{V}\mathbf{X}}\mathbf{U}]\mathbf{y}$. Correspondingly, the WLSE of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$ is defined to be

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) := \mathbf{X}\tilde{\boldsymbol{\beta}} = [\mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V} + \mathbf{X}\mathbf{F}_{\mathbf{V}\mathbf{X}}\mathbf{U}]\mathbf{y}. \tag{1.8}$$

Further, let $\mathbf{P}_{\mathbf{X};\mathbf{V}}$ and $\mathbf{P}_{\mathbf{X}:\mathbf{V}}$ denote

$$\mathbf{P}_{\mathbf{X};\mathbf{V}} = \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}, \quad \mathbf{P}_{\mathbf{X}:\mathbf{V}} = \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V} + \mathbf{X}\mathbf{F}_{\mathbf{V}\mathbf{X}}\mathbf{U} = \mathbf{P}_{\mathbf{X};\mathbf{V}} + \mathbf{X}\mathbf{F}_{\mathbf{V}\mathbf{X}}\mathbf{U}, \tag{1.9}$$

both of which are called projectors into $\mathscr{R}(\mathbf{X})$ with respect to the seminorm $\|\cdot\|_{\mathbf{V}}$, see Rao and Mitra (1971a,b) and Mitra and Rao (1974). In what follows, we take the homogeneous estimator in (1.8) as the general expression of WLSEs of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$. In addition, we use $\{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\}$ to denote the collection of all $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$.

It can be seen from (1.7) and (1.8) that for a given weight matrix $\mathbf{V}$, the two estimators $\tilde{\boldsymbol{\beta}}$ and $\mathbf{X}\tilde{\boldsymbol{\beta}}$ are not necessarily unbiased for $\boldsymbol{\beta}$ and $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$. However, it is easy to show that for any given weight matrix $\mathbf{V}$, there exists a matrix $\mathbf{U}$ such that $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ in (1.8) is unbiased for $\mathbf{X}\boldsymbol{\beta}$.

According to (1.8), the WLSEs of $\mathbf{X}_1\boldsymbol{\beta}_1$ and $\mathbf{X}_2\boldsymbol{\beta}_2$ under the two small models in (1.3) can be written as

$$\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) = \mathbf{P}_{\mathbf{X}_1:\mathbf{V}}\mathbf{y}, \quad \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2) = \mathbf{P}_{\mathbf{X}_2:\mathbf{V}}\mathbf{y}, \tag{1.10}$$

where

$$\mathbf{P}_{\mathbf{X}_i:\mathbf{V}} = \mathbf{P}_{\mathbf{X}_i;\mathbf{V}} + \mathbf{X}_i\mathbf{F}_{\mathbf{V}\mathbf{X}_i}\mathbf{U}_i = \mathbf{X}_i(\mathbf{X}_i'\mathbf{V}\mathbf{X}_i)^{+}\mathbf{X}_i'\mathbf{V} + \mathbf{X}_i\mathbf{F}_{\mathbf{V}\mathbf{X}_i}\mathbf{U}_i, \quad i = 1, 2,$$

and $\mathbf{U}_1 \in \mathbb{R}^{p_1\times n}$ and $\mathbf{U}_2 \in \mathbb{R}^{p_2\times n}$ are arbitrary.

Because there are arbitrary matrices $\mathbf{U}, \mathbf{U}_1$, and $\mathbf{U}_2$ in (1.8) and (1.10), it is possible to take $\mathbf{U}, \mathbf{U}_1$, and $\mathbf{U}_2$ such that the WLSEs have some prescribed properties, such as, unbiasedness, minimum norm, minimum covariance, etc. In statistical applications, the weight matrix $\mathbf{V}$ is often taken as $\mathbf{V} = \Sigma^{-}$ or $\mathbf{V} = (\mathbf{X}\mathbf{T}\mathbf{X}' + \Sigma)^{-}$, where $\mathbf{T}$ is a non-negative definite matrix such that $r(\mathbf{X}\mathbf{T}\mathbf{X}' + \Sigma) = r[\mathbf{X},\ \Sigma]$. In particular, if Σ is positive definite and $\mathbf{X}$ has full column rank, then

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y} \tag{1.11}$$

is the well-known unique best unbiased linear estimator (BLUE) of $\mathbf{X}\boldsymbol{\beta}$ under $\mathscr{M}$.

Notice that the WLSEs in (1.8) and (1.10) are not necessarily unique. Extensions of (1.5) to the WLSEs under $\mathscr{M}$, $\mathscr{M}_1$, and $\mathscr{M}_2$ have the following three cases:

(I) $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ holds for some $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$, and $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$;
(II) $\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} \subseteq \{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\}$;
(III) $\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} = \{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\}$.

In Sec. 2, we give a variety of necessary and sufficient conditions for the three assertions to hold. As consequences, we give a group of necessary and sufficient conditions for the BLUE in (1.11) to be the sum of two WLSEs under $\mathscr{M}_1$ and $\mathscr{M}_2$ in (1.10). The proofs of main results are given in Appendix.

Because the WLSEs in (1.8) and (1.10) are matrix pencils consisting of Moore–Penrose inverses and arbitrary matrices, we need to use the following rank formulas for partitioned matrices due to Marsaglia and Styan (1974) to simplify various matrix operations related to the WLSEs.

Lemma 1.2. *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*

$$r[\mathbf{A},\ \mathbf{B}] = r(\mathbf{A}) + r(\mathbf{E}_{\mathbf{A}}\mathbf{B}) = r(\mathbf{B}) + r(\mathbf{E}_{\mathbf{B}}\mathbf{A}), \tag{1.12}$$

$$r\begin{bmatrix}\mathbf{A}\\ \mathbf{C}\end{bmatrix} = r(\mathbf{A}) + r(\mathbf{C}\mathbf{F}_{\mathbf{A}}) = r(\mathbf{C}) + r(\mathbf{A}\mathbf{F}_{\mathbf{C}}). \tag{1.13}$$

It is easy to see $r(\mathbf{B} - \mathbf{A}\mathbf{A}^{+}\mathbf{B}) \geq r(\mathbf{B}) - r(\mathbf{A}\mathbf{A}^{+}\mathbf{B}) = r(\mathbf{B}) - r(\mathbf{A}'\mathbf{B})$. Therefore, it follows from (1.12) that

$$r[\mathbf{A},\ \mathbf{B}] \geq r(\mathbf{A}) + r(\mathbf{B}) - r(\mathbf{A}'\mathbf{B}). \tag{1.14}$$

Moreover, we can find by (1.12) the following two results:

$$\mathscr{R}(\mathbf{B}) \subseteq \mathscr{R}(\mathbf{A}) \Leftrightarrow \mathbf{A}\mathbf{A}^{+}\mathbf{B} = \mathbf{B} \Leftrightarrow r[\mathbf{A},\ \mathbf{B}] = r(\mathbf{A}), \tag{1.15}$$

$$\mathscr{R}(\mathbf{A}_1) = \mathscr{R}(\mathbf{A}_2) \quad \text{and} \quad \mathscr{R}(\mathbf{B}_1) = \mathscr{R}(\mathbf{B}_2) \Rightarrow r[\mathbf{A}_1,\ \mathbf{B}_1] = r[\mathbf{A}_2,\ \mathbf{B}_2]. \tag{1.16}$$

Lemma 1.3. *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$, *and let* $\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3 \in \mathbb{R}^{n\times m}$ *be three outer inverses of* $\mathbf{A}$, *i.e.,* $\mathbf{Z}_i\mathbf{A}\mathbf{Z}_i = \mathbf{Z}_i$, $i = 1, 2, 3$. *Also suppose* $\mathscr{R}(\mathbf{Z}_i) \subseteq \mathscr{R}(\mathbf{Z}_1)$ *and* $\mathscr{R}(\mathbf{Z}_i') \subseteq \mathscr{R}(\mathbf{Z}_1')$, $i = 2, 3$. *Then*

$$r(\mathbf{Z}_1 - \mathbf{Z}_2 - \mathbf{Z}_3) = r(\mathbf{Z}_1) - r(\mathbf{Z}_2) - r(\mathbf{Z}_3) + r(\mathbf{Z}_2\mathbf{A}\mathbf{Z}_3) + r(\mathbf{Z}_3\mathbf{A}\mathbf{Z}_2). \tag{1.17}$$

The following results are shown in Tian (2002) and Tian and Cheng (2003).

Lemma 1.4. *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times k}$ *and* $\mathbf{C} \in \mathbb{R}^{l\times n}$ *be given. Then*

$$\max_{\mathbf{Z}\in\mathbb{R}^{k\times l}} r(\mathbf{A} - \mathbf{B}\mathbf{Z}\mathbf{C}) = \min\left\{ r[\mathbf{A},\ \mathbf{B}],\ r\begin{bmatrix}\mathbf{A}\\ \mathbf{C}\end{bmatrix} \right\}, \tag{1.18}$$

$$\min_{\mathbf{Z}\in\mathbb{R}^{k\times l}} r(\mathbf{A} - \mathbf{B}\mathbf{Z}\mathbf{C}) = r[\mathbf{A},\ \mathbf{B}] + r\begin{bmatrix}\mathbf{A}\\ \mathbf{C}\end{bmatrix} - r\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{0}\end{bmatrix}. \tag{1.19}$$

In particular,

$$\mathbf{B}\mathbf{Z}\mathbf{C} = \mathbf{A} \textit{ is consistent} \Leftrightarrow r[\mathbf{A},\ \mathbf{B}] = r(\mathbf{B}) \quad \text{and} \quad r\begin{bmatrix}\mathbf{A}\\ \mathbf{C}\end{bmatrix} = r(\mathbf{C}). \tag{1.20}$$

2. Sum Decompositions of WLSEs

In order to characterize the sum decompositions of the WLSEs described in Sec. 1, we assume in what follows that the model $\mathscr{M}$ in (1.2) is correct. In this case, the two small models $\mathscr{M}_1$ and $\mathscr{M}_2$ in (1.3) are in fact two misspecified models of $\mathscr{M}$. It is easy to show that if $\mathscr{M}$ in (1.2) is correct, then

$$\mathbf{y} \in \mathscr{R}[\mathbf{X},\ \Sigma] \tag{2.1}$$

holds with probability 1, see Rao (1971, 1973). Hence, (2.1) should be taken into account when investigating various properties of estimators under $\mathscr{M}$ in (1.2). A linear model is said to be consistent if it satisfies (2.1). In this case, a pair of linear estimators $\mathbf{L}_1\mathbf{y}$ and $\mathbf{L}_2\mathbf{y}$ are said to be equal with probability 1 under the consistent model $\mathscr{M}$ if

$$(\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X},\ \Sigma] = \mathbf{0}. \tag{2.2}$$

If the model $\mathscr{M}$ in (1.2) is correct, then it is consistent, too. However, the consistency of a model does not imply that it is correct. In fact, if Σ is positive definite, then the correct model $\mathscr{M}$ in (1.2), as well as the two small (misspecified) models $\mathscr{M}_1$ and $\mathscr{M}_2$ in (1.3) are always consistent. If $r[\mathbf{X},\ \Sigma] < n$, the consistency of $\mathscr{M}$ in (1.2) does not guarantee the consistency of the two models in (1.3). Because $\mathscr{M}_1$ and $\mathscr{M}_2$ are two misspecified models of $\mathscr{M}$, we cannot assume that $\mathbf{y} \in \mathscr{R}[\mathbf{X}_1,\ \Sigma]$ and $\mathbf{y} \in \mathscr{R}[\mathbf{X}_2, \Sigma]$. Instead, we assume that the vector $\mathbf{y}$ in $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ and $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ only satisfies (2.1).

Two main results of the present article on sum decomposition of $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ are given below.

Theorem 2.1. *Let* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$, *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *be as given in* (1.8) *and* (1.10). *Then the following statements are equivalent*:

(a) *There exist* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$, *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *such that*

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2) \tag{2.3}$$

holds with probability 1.

(b) *The set inclusion* $\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} \subseteq \{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\}$ *holds with probability* 1.

(c) $\mathbf{VP}_{\mathbf{X};\mathbf{V}} = \mathbf{VP}_{\mathbf{X}_1;\mathbf{V}} + \mathbf{VP}_{\mathbf{X}_2;\mathbf{V}}$.

(d) $\mathbf{X}_1'\mathbf{V}\mathbf{X}_2 = \mathbf{0}$, *i.e.*, $\mathbf{X}_1$ *and* $\mathbf{X}_2$ *are* $\mathbf{V}$*-orthogonal.*

Theorem 2.2. *Let* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$, *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *be as given in* (1.8) *and* (1.10). *Then the following statements are equivalent*:

(a) *The set inclusion* $\{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\} \subseteq \{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\}$ *holds with probability* 1.

(b) $r(\mathbf{X}) + 2r(\mathbf{X}_1'\mathbf{V}\mathbf{X}_2) = r(\mathbf{N})$, *where* $\mathbf{N} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{V}\mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V}\mathbf{X}_2 \end{bmatrix}$.

Combining Theorems 2.1 and 2.2 gives the following result.

Theorem 2.3. *Let* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *be as given in* (1.8) *and* (1.10). *Then the following statements are equivalent*:

(a) *The set equality* $\{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\} = \{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\}$ *holds with probability* 1.

(b) $\mathbf{X}_1'\mathbf{V}\mathbf{X}_2 = \mathbf{0}$ *and* $\mathscr{R}\begin{bmatrix} \mathbf{X}_1'\mathbf{V} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2'\mathbf{V} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{bmatrix}$.

The results in Theorem 2.1 (a), (b), and (c) can be regarded as extensions of (1.4) and (1.5), while $\mathbf{X}_1'\mathbf{V}\mathbf{X}_2 = \mathbf{0}$ is an extension of the orthogonal equality $\mathbf{X}_1'\mathbf{X}_2 = \mathbf{0}$. Under the conditions in the previous theorems, the sum decompositions can be used to reduce the computation of $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and to derive statistical properties of $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$.

Because $\mathscr{M}$ in (1.2) is a general linear model and the matrix Σ occurs in (2.2), it is somehow surprising that the necessary and sufficient conditions for the sum decompositions of $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ in Theorems 2.1, 2.2, and 2.3 only consist of the model matrix $\mathbf{X}$ and the weight matrix $\mathbf{V}$. It should be pointed out that the equality in (2.3) does not imply that $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ and $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ are uncorrelated. In fact, it is easy to derive from (1.10) that the correlation matrix between $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ and $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ is

$$Cov\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1),\ \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} = \sigma^2\mathbf{P}_{\mathbf{X}_1:\mathbf{V}}\Sigma\mathbf{P}_{\mathbf{X}_2:\mathbf{V}}'. \tag{2.4}$$

This is a quadratic form with respect to the two arbitrary matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ in $\mathbf{P}_{\mathbf{X}_1:\mathbf{V}}$ and $\mathbf{P}_{\mathbf{X}_2:\mathbf{V}}$. Also note that the covariance matrix $\sigma^2\Sigma$ occurs in (2.4), so that it is a challenging problem to give necessary and sufficient conditions for (2.4) to be null.

Concerning the uniqueness of the estimators in (1.8) and (1.10), as well as the sum decomposition of $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ when it is unique, we have the following two theorems.

Theorem 2.4. *Let* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$, $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$, *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *be as given in* (1.8) *and* (1.10). *Then*:

(a) $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *is unique if and only if* $r(\mathbf{V}\mathbf{X}) = r(\mathbf{X})$, *i.e.*, $\mathscr{R}(\mathbf{X}'\mathbf{V}) = \mathscr{R}(\mathbf{X}')$. *In this case, the unique* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X};\mathbf{V}}\mathbf{y} = \mathbf{X}(\mathbf{X}'\mathbf{V}\mathbf{X})^{+}\mathbf{X}'\mathbf{V}\mathbf{y}$ *is unbiased for* $\mathbf{X}\boldsymbol{\beta}$.

(b) $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *are unique if and only if* $r(\mathbf{V}\mathbf{X}_1) = r(\mathbf{X}_1)$ *and* $r(\mathbf{V}\mathbf{X}_2) = r(\mathbf{X}_2)$ *hold, respectively.*

Theorem 2.5. *Suppose that* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *in* (1.8) *is unique. Then*:

(a) *Both* $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *in* (1.10) *are unique.*

(b) $Cov\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1),\ \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} = \sigma^2\mathbf{X}_1(\mathbf{X}_1'\mathbf{V}\mathbf{X}_1)^{+}\mathbf{X}_1'\mathbf{V}\Sigma\mathbf{V}\mathbf{X}_2(\mathbf{X}_2'\mathbf{V}\mathbf{X}_2)^{+}\mathbf{X}_2'$.

(c) $Cov\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1),\ \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} = \mathbf{0}$ *if and only if* $\mathbf{X}_1'\mathbf{V}\Sigma\mathbf{V}\mathbf{X}_2 = \mathbf{0}$.

(d) *The sum decomposition* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *holds with probability* 1 *if and only if* $\mathbf{P}_{\mathbf{X};\mathbf{V}} = \mathbf{P}_{\mathbf{X}_1;\mathbf{V}} + \mathbf{P}_{\mathbf{X}_2;\mathbf{V}}$, *or equivalently,* $\mathbf{X}_1'\mathbf{V}\mathbf{X}_2 = \mathbf{0}$.

As mentioned in Sec. 1, the weight matrix $\mathbf{V}$ in (1.6) is often taken as $\mathbf{V} = \boldsymbol{\Sigma}^{-}$ or $\mathbf{V} = (\mathbf{X}\mathbf{T}\mathbf{X}' + \boldsymbol{\Sigma})^{-}$ with $r(\mathbf{X}\mathbf{T}\mathbf{X}' + \boldsymbol{\Sigma}) = r[\mathbf{X},\ \boldsymbol{\Sigma}]$. In this case, the previous results can be simplified further. In particular, if $\boldsymbol{\Sigma}$ is positive definite and $r(\mathbf{X}) = p$ in (1.2), and the weight matrix $\mathbf{V}$ is taken as $\mathbf{V} = \boldsymbol{\Sigma}^{-1}$ in (1.8) and (1.10), then we have the following consequences.

Lemma 2.1. *Suppose* $\boldsymbol{\Sigma}$ *is positive definite and* $r(\mathbf{X}) = p$ *in* (1.2), *and let* $\mathbf{V} = \Sigma^{-1}$ *in* (1.8) *and* (1.10). *Then*:

(a) *The unique BLUE of* $\mathbf{X}\boldsymbol{\beta}$ *under* $\mathscr{M}$ *is*

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y} \tag{2.5}$$

with $E[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta}$ *and* $Cov[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \sigma^2\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'$.

(b) *The unique WLSEs of* $\mathbf{X}_i\boldsymbol{\beta}_i$ *under* $\mathscr{M}_i$ *are*

$$\mathrm{WLSE}_{\mathscr{M}_i}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathbf{X}_i(\mathbf{X}_i'\Sigma^{-1}\mathbf{X}_i)^{-1}\mathbf{X}_i'\Sigma^{-1}\mathbf{y}, \quad i = 1, 2 \tag{2.6}$$

with

$$E[\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)] = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_1(\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_2\boldsymbol{\beta}_2, \tag{2.7}$$

$$E[\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)] = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{X}_2(\mathbf{X}_2'\Sigma^{-1}\mathbf{X}_2)^{-1}\mathbf{X}_2'\Sigma^{-1}\mathbf{X}_1\boldsymbol{\beta}_1, \tag{2.8}$$

$$Cov[\mathrm{WLSE}_{\mathscr{M}_i}(\mathbf{X}_i\boldsymbol{\beta}_i)] = \sigma^2\mathbf{X}_i(\mathbf{X}_i'\Sigma^{-1}\mathbf{X}_i)^{-1}\mathbf{X}_i', \quad i = 1, 2. \tag{2.9}$$

(c) *The covariance matrix between* $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *and* $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$ *is*

$$\begin{aligned} &Cov\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1),\ \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} \\ &\quad = \sigma^2\mathbf{X}_1(\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_2(\mathbf{X}_2'\Sigma^{-1}\mathbf{X}_2)^{-1}\mathbf{X}_2'. \end{aligned} \tag{2.10}$$

Applying Theorems 2.1 and 2.4 to (2.5)–(2.10) gives the following result.

Corollary 2.1. *Let* $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *and* $\mathrm{WLSE}_{\mathscr{M}_i}(\mathbf{X}_i\boldsymbol{\beta}_i)$ *be as given in* (2.5) *and* (2.6). *Then the following statements are equivalent*:

(a) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$.
(b) $E[\mathrm{WLSE}_{\mathscr{M}_i}(\mathbf{X}_i\boldsymbol{\beta}_i)] = \mathbf{X}_i\boldsymbol{\beta}_i$, $i = 1, 2$.
(c) $Cov\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1),\ \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)\} = \mathbf{0}$.
(d) $Cov[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = Cov[\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)] + Cov[\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)]$.
(e) $\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_2 = \mathbf{0}$.

Finally, we give two results on relationships between $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ and $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$.

Theorem 2.6. *Let* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *and* $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *be as given in* (1.8) *and* (1.10). *Then the following statements are equivalent*:

(a) *There exist* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *and* $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *such that the equality*

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) \tag{2.11}$$

holds with probability 1.
(b) *The set inclusion* $\{\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)\} \subseteq \{\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\}$ *holds with probability* 1.
(c) $\mathbf{VP}_{\mathbf{X}:\mathbf{V}} = \mathbf{VP}_{\mathbf{X}_1:\mathbf{V}}$.
(d) $\mathscr{R}(\mathbf{VX}_2) \subseteq \mathscr{R}(\mathbf{VX}_1)$.

The following result is a direct consequence of Theorem 2.6.

Corollary 2.2. *Suppose* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ *in* (1.8) *is unique. Then the equality* $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ *holds with probability* 1 *if and only if* $\mathscr{R}(\mathbf{X}_2) \subseteq \mathscr{R}(\mathbf{X}_1)$.

Remark 2.1. In this article, we have obtained a variety of necessary and sufficient conditions for WLSEs of $\mathbf{X}\boldsymbol{\beta}$ under (1.2) to be sums of WLSEs under (1.3). Under these conditions, it is expected that the sum decompositions can be used to derive some valuable statistical properties of WLSEs and BLUEs of $\mathbf{X}\boldsymbol{\beta}$ under (1.2).

The results in this article can also be extended to some more general settings. Two future research topics on sum decompositions of WLSEs under partitioned linear models are given below:

(a) For the general partitioned linear model

$$\mathscr{M} = \left\{\mathbf{y},\ \mathbf{X}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{X}_k\boldsymbol{\beta}_k,\ \sigma^2\Sigma\right\}$$

and its k small models $\mathscr{M}_i = \{\mathbf{y},\ \mathbf{X}_i\boldsymbol{\beta}_i,\ \sigma^2\Sigma\}$, $i = 1, \ldots, k$, establish necessary and sufficient conditions for the sum decomposition

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{X}_k\boldsymbol{\beta}_k) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) + \cdots + \mathrm{WLSE}_{\mathscr{M}_k}(\mathbf{X}_k\boldsymbol{\beta}_k)$$

to hold.

(b) Suppose $\mathbf{K} = [\mathbf{K}_1,\ \mathbf{K}_2] \in \mathbb{R}^{q\times(p_1+p_2)}$ is given and $\mathbf{K}\boldsymbol{\beta} = \mathbf{K}_1\boldsymbol{\beta}_1 + \mathbf{K}_2\boldsymbol{\beta}_2$ is estimable under $\mathscr{M}$ in (1.2), i.e., $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$. Then establish necessary and sufficient conditions for the sum decomposition

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{K}_1\boldsymbol{\beta}_1) + \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{K}_2\boldsymbol{\beta}_2)$$

to hold.

Appendix

Recall that the rank of a matrix is defined to the dimension of the row or column space of the matrix. Also recall that $\mathbf{A} = \mathbf{0}$ if and only if $r(\mathbf{A}) = 0$. From this simple fact we see that two matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size are equal if and only if $r(\mathbf{A} - \mathbf{B}) = 0$; two sets S_1 and S_2 consisting of matrices of the same size have a common matrix if and only if

$$\min_{\mathbf{A}\in S_1,\ \mathbf{B}\in S_2} r(\mathbf{A} - \mathbf{B}) = 0;$$

the set inclusion $S_1 \subseteq S_2$ if and only if

$$\max_{\mathbf{A}\in S_1} \min_{\mathbf{B}\in S_2} r(\mathbf{A} - \mathbf{B}) = 0.$$

If $\mathbf{A} - \mathbf{B}$ can be written as a linear matrix expression with some arbitrary matrices, then we can find the extremal ranks of this expression by (1.18) and (1.19), and use the extremal ranks to characterize relations between the two sets S_1 and S_2. This method is available for studying various matrix expressions involving arbitrary matrices. In Puntanen et al. (2005), Qian and Tian (2006), and Tian and Wiens

(2006), the matrix rank method is widely used to characterize a variety of equalities for estimators under $\mathscr{M}$. In the Appendix, we also use the method to prove the results in Sec. 2.

Proof of Lemma 1.3. Notice that the rank of a matrix is invariant under elementary block matrix operations. Hence, it is easy to find by elementary block matrix operations that

$$\begin{aligned} r\begin{bmatrix} -\mathbf{Z}_1 & \mathbf{0} & \mathbf{0} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{Z}_2 & \mathbf{0} & \mathbf{Z}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_3 & \mathbf{Z}_3 \\ \mathbf{Z}_1 & \mathbf{Z}_2 & \mathbf{Z}_3 & \mathbf{0} \end{bmatrix} &= r\begin{bmatrix} -\mathbf{Z}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Z}_1 - \mathbf{Z}_2 - \mathbf{Z}_3 \end{bmatrix} \\ &= r(\mathbf{Z}_1 - \mathbf{Z}_2 - \mathbf{Z}_3) + r(\mathbf{Z}_1) + r(\mathbf{Z}_2) + r(\mathbf{Z}_3). \end{aligned} \tag{3.1}$$

Under the given conditions in Lemma 1.3, we also find by elementary block matrix operations that

$$\begin{aligned} r\begin{bmatrix} -\mathbf{Z}_1 & \mathbf{0} & \mathbf{0} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{Z}_2 & \mathbf{0} & \mathbf{Z}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_3 & \mathbf{Z}_3 \\ \mathbf{Z}_1 & \mathbf{Z}_2 & \mathbf{Z}_3 & \mathbf{0} \end{bmatrix} &= r\begin{bmatrix} \mathbf{0} & \mathbf{Z}_1\mathbf{A}\mathbf{Z}_2 & \mathbf{Z}_1\mathbf{A}\mathbf{Z}_3 & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{Z}_2 & \mathbf{0} & \mathbf{Z}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{Z}_3 & \mathbf{Z}_3 \\ \mathbf{Z}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \\ &= r\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{0} & -\mathbf{Z}_2\mathbf{A}\mathbf{Z}_3 & \mathbf{0} \\ \mathbf{0} & -\mathbf{Z}_3\mathbf{A}\mathbf{Z}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{Z}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \\ &= 2r(\mathbf{Z}_1) + r(\mathbf{Z}_2\mathbf{A}\mathbf{Z}_3) + r(\mathbf{Z}_3\mathbf{A}\mathbf{Z}_2). \end{aligned} \tag{3.2}$$

Combining (3.1) and (3.2) results in (1.17). □

Proof of Theorem 2.1. We first show that the following two results

$$\mathscr{R}[\mathbf{X}_1\mathbf{F}_{\mathbf{V}\mathbf{X}_1}, \ \mathbf{X}_2\mathbf{F}_{\mathbf{V}\mathbf{X}_2}] \subseteq \mathscr{R}(\mathbf{X}\mathbf{F}_{\mathbf{V}\mathbf{X}}), \tag{3.3}$$

$$r(\mathbf{V}\mathbf{X}) \geq r(\mathbf{V}\mathbf{X}_1) + r(\mathbf{V}\mathbf{X}_2) - r(\mathbf{X}_1'\mathbf{V}\mathbf{X}_2) \tag{3.4}$$

hold. Applying (1.13) to the matrices in (3.3) and simplifying by elementary block matrix operations yield

$$r(\mathbf{X}\mathbf{F}_{\mathbf{V}\mathbf{X}}) = r\begin{bmatrix} \mathbf{X} \\ \mathbf{V}\mathbf{X} \end{bmatrix} - r(\mathbf{V}\mathbf{X}) = r(\mathbf{X}) - r(\mathbf{V}\mathbf{X}),$$

$$\begin{aligned} r[\mathbf{X}\mathbf{F}_{\mathbf{V}\mathbf{X}}, \ \mathbf{X}_1\mathbf{F}_{\mathbf{V}\mathbf{X}_1}, \ \mathbf{X}_2\mathbf{F}_{\mathbf{V}}\mathbf{X}_2] &= r\begin{bmatrix} \mathbf{X} & \mathbf{X}_1 & \mathbf{X}_2 \\ \mathbf{V}\mathbf{X} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}\mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}\mathbf{X}_2 \end{bmatrix} - r(\mathbf{V}\mathbf{X}) - r(\mathbf{V}\mathbf{X}_1) - r(\mathbf{V}\mathbf{X}_2) \\ &= r\begin{bmatrix} \mathbf{X} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{V}\mathbf{X}_1 & -\mathbf{V}\mathbf{X}_2 \\ \mathbf{0} & \mathbf{V}\mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}\mathbf{X}_2 \end{bmatrix} - r(\mathbf{V}\mathbf{X}) - r(\mathbf{V}\mathbf{X}_1) - r(\mathbf{V}\mathbf{X}_2) \end{aligned}$$

$$= r\begin{bmatrix} \mathbf{X} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{VX}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{VX}_2 \end{bmatrix} - r(\mathbf{VX}) - r(\mathbf{VX}_1) - r(\mathbf{VX}_2)$$
$$= r(\mathbf{X}) - r(\mathbf{VX}).$$

Hence,

$$r[\mathbf{XF}_{\mathbf{VX}},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}] = r(\mathbf{XF}_{\mathbf{VX}}) = r(\mathbf{X}) - r(\mathbf{VX}), \tag{3.5}$$

which implies (3.3) by (1.15). Applying (1.14) to $\mathbf{VX}$ gives

$$\begin{aligned} r(\mathbf{VX}) = r(\mathbf{V}^{1/2}\mathbf{X}) &= r[\mathbf{V}^{1/2}\mathbf{X}_1,\ \mathbf{V}^{1/2}\mathbf{X}_2] \\ &\geq r(\mathbf{V}^{1/2}\mathbf{X}_1) + r(\mathbf{V}^{1/2}\mathbf{X}_2) - r(\mathbf{X}_1'\mathbf{VX}_2) \\ &= r(\mathbf{VX}_1) + r(\mathbf{VX}_2) - r(\mathbf{X}_1'\mathbf{VX}_2), \end{aligned}$$

establishing (3.4), where $\mathbf{V}^{1/2}$ is the square root of the nonnegative definite matrix $\mathbf{V}$.

It follows from (1.6) and (1.10) that

$$\begin{aligned} &\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) - \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) - \mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2) \\ &\quad = \mathbf{P}_{\mathbf{X}:\mathbf{V}}\mathbf{y} - \mathbf{P}_{\mathbf{X}_1:\mathbf{V}}\mathbf{y} - \mathbf{P}_{\mathbf{X}_2:\mathbf{V}}\mathbf{y} \\ &\quad = (\mathbf{GV} + \mathbf{XF}_{\mathbf{VX}}\mathbf{U} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1 - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2)\mathbf{y}, \end{aligned}$$

where $\mathbf{G} = \mathbf{X}(\mathbf{X}'\mathbf{VX})^+\mathbf{X}' - \mathbf{X}_1(\mathbf{X}_1'\mathbf{VX}_1)^+\mathbf{X}_1' - \mathbf{X}_2(\mathbf{X}_2'\mathbf{VX}_2)^+\mathbf{X}_2'$. Hence it can be seen from (2.2) that (2.3) holds with probability 1 if and only if there exist $\mathbf{U}$, $\mathbf{U}_1$, and $\mathbf{U}_2$ such that

$$(\mathbf{GV} + \mathbf{XF}_{\mathbf{VX}}\mathbf{U} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1 - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2)\mathbf{y} = \mathbf{0} \quad \text{for all } \mathbf{y} \in \mathscr{R}[\mathbf{X},\ \Sigma],$$

that is, there exist $\mathbf{U}$, $\mathbf{U}_1$, and $\mathbf{U}_2$ such that

$$(\mathbf{GV} + \mathbf{XF}_{\mathbf{VX}}\mathbf{U} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1 - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2)\mathbf{S} = \mathbf{0}, \tag{3.6}$$

where $\mathbf{S} = [\mathbf{X},\ \Sigma]$. Rewrite (3.6) as

$$\mathbf{AZS} = -\mathbf{GVS}, \tag{3.7}$$

where $\mathbf{A} = [\mathbf{XF}_{\mathbf{VX}},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}]$ and $\mathbf{Z} = [\mathbf{U}', -\mathbf{U}_1', -\mathbf{U}_2']'$. From (1.20), the equation in (3.7) is solvable for $\mathbf{Z}$ if and only if

$$r[\mathbf{GVS},\ \mathbf{A}] = r(\mathbf{A}) \quad \text{and} \quad r\begin{bmatrix} \mathbf{GVS} \\ \mathbf{S} \end{bmatrix} = r(\mathbf{S}). \tag{3.8}$$

The second equality in (3.8) holds naturally. It is easy to see $\mathscr{R}(\mathbf{G}) = \mathscr{R}(\mathbf{G}') \subseteq \mathscr{R}(\mathbf{S})$, i.e., $\mathbf{SS}^+\mathbf{G} = \mathbf{G}$. Hence it follows that $\mathscr{R}(\mathbf{GVS}) \supseteq \mathscr{R}(\mathbf{GVSS}^+\mathbf{G}) = \mathscr{R}(\mathbf{GVG}) = \mathscr{R}(\mathbf{GV})$, which obviously implies that

$$\mathscr{R}(\mathbf{GVS}) = \mathscr{R}(\mathbf{GV}). \tag{3.9}$$

Under (3.3) and (3.9), it can be shown from (1.16) and (3.5) that the first equality in (3.8) is equivalent to

$$r[\mathbf{GV},\ \mathbf{XF}_{\mathbf{VX}}] = r(\mathbf{XF}_{\mathbf{VX}}) = r(\mathbf{X}) - r(\mathbf{VX}). \tag{3.10}$$

Applying (1.13) to the left-hand side of (3.10) and simplifying by elementary block matrix operations yields:

$$\begin{aligned} r[\mathbf{GV},\ \mathbf{XF}_{\mathbf{VX}}] &= r\begin{bmatrix} \mathbf{GV} & \mathbf{X} \\ \mathbf{0} & \mathbf{VX} \end{bmatrix} - r(\mathbf{VX}) = r\begin{bmatrix} \mathbf{0} & \mathbf{X} \\ -\mathbf{VGV} & \mathbf{0} \end{bmatrix} - r(\mathbf{VX}) \\ &= r(\mathbf{VGV}) + r(\mathbf{X}) - r(\mathbf{VX}). \end{aligned} \tag{3.11}$$

Hence, (3.10) is equivalent to $\mathbf{VGV} = \mathbf{0}$, which is the equality in (c). Let

$$\mathbf{Z} = \mathbf{VX}(\mathbf{X}'\mathbf{VX})^{+}\mathbf{X}'\mathbf{V}, \quad \mathbf{Z}_1 = \mathbf{VX}_1(\mathbf{X}_1'\mathbf{VX}_1)^{+}\mathbf{X}_1'\mathbf{V}, \quad \mathbf{Z}_2 = \mathbf{VX}_2(\mathbf{X}_2'\mathbf{VX}_2)^{+}\mathbf{X}_2'\mathbf{V}.$$

Then it is easy to verify that the nonnegative definite matrices $\mathbf{Z}$, $\mathbf{Z}_1$, and $\mathbf{Z}_2$ are outer inverses of $\mathbf{V}^{+}$, and

$$\mathscr{R}(\mathbf{Z}) = \mathscr{R}(\mathbf{VX}), \quad \mathscr{R}(\mathbf{Z}_i) = \mathscr{R}(\mathbf{VX}_i), \quad \mathscr{R}(\mathbf{VX}_i) \subseteq \mathscr{R}(\mathbf{VX}), \quad \mathbf{Z}_i\mathbf{X}_i = \mathbf{VX}_i, \quad i = 1, 2 \tag{3.12}$$

hold. Under these conditions, applying (1.17) to $\mathbf{VGV}$ gives

$$\begin{aligned} r(\mathbf{VGV}) &= r(\mathbf{Z} - \mathbf{Z}_1 - \mathbf{Z}_2) \\ &= r(\mathbf{Z}) - r(\mathbf{Z}_1) - r(\mathbf{Z}_2) + 2r(\mathbf{Z}_1\mathbf{V}^{+}\mathbf{Z}_2) \\ &= r(\mathbf{VX}) - r(\mathbf{VX}_1) - r(\mathbf{VX}_2) + 2r(\mathbf{X}_1'\mathbf{VX}_2) \\ &= [r(\mathbf{VX}) + r(\mathbf{X}_1'\mathbf{VX}_2) - r(\mathbf{VX}_1) - r(\mathbf{VX}_2)] + r(\mathbf{X}_1'\mathbf{VX}_2). \end{aligned} \tag{3.13}$$

The equivalence of $\mathbf{VGV} = \mathbf{0}$ and $\mathbf{X}_1'\mathbf{VX}_2 = \mathbf{0}$ follows from (3.4) and (3.13).

It can also be seen from (3.6) that the set inclusion in Theorem 2.1(b) holds with probability 1 if and only if

$$\min_{\mathbf{U}} r(\mathbf{GVS} + \mathbf{XF}_{\mathbf{VX}}\mathbf{US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S} - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2\mathbf{S}) = 0$$

holds for any $\mathbf{U}_1$ and $\mathbf{U}_2$, where $\mathbf{S} = [\mathbf{X},\ \Sigma]$. Under (3.3) and (3.9), applying (1.19) gives

$$\begin{aligned} &\min_{\mathbf{U}} r(\mathbf{GVS} + \mathbf{XF}_{\mathbf{VX}}\mathbf{US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S} - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2\mathbf{S}) \\ &\quad = r[\mathbf{GVS} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S} - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2\mathbf{S},\ \mathbf{XF}_{\mathbf{VX}}] - r(\mathbf{XF}_{\mathbf{VX}}) \\ &\quad = r[\mathbf{GV},\ \mathbf{XF}_{\mathbf{VX}}] - r(\mathbf{XF}_{\mathbf{VX}}) \\ &\quad = 2r(\mathbf{X}_1'\mathbf{VX}_2) + r(\mathbf{VX}) - r(\mathbf{VX}_1) - r(\mathbf{VX}_2) \quad \text{(by (3.5), (3.11) and (3.13)).} \end{aligned}$$

Hence, (b) is equivalent to (d) as well. □

Proof of Theorem 2.2. It can be seen from (3.6) that the set inclusion in (a) holds if and only if there exist $\mathbf{U}$, $\mathbf{U}_1$, and $\mathbf{U}_2$ such that

$$\max_{\mathbf{U}} \min_{\mathbf{U}_1,\mathbf{U}_2} r(\mathbf{GVS} + \mathbf{XF_{VX}US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S} - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2\mathbf{S}) = 0. \tag{3.14}$$

It follows from (1.19) that

$$\begin{aligned}
&\min_{\mathbf{U}_1,\mathbf{U}_2} r(\mathbf{GVS} + \mathbf{XF_{VX}US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S} - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2\mathbf{S}) \\
&\quad = \min_{\mathbf{U}_1,\mathbf{U}_2} r\left(\mathbf{GVS} + \mathbf{XF_{VX}US} - [\mathbf{XF_{VX}},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}]\begin{bmatrix}\mathbf{U}_1\\ \mathbf{U}_2\end{bmatrix}\mathbf{S}\right) \\
&\quad = r[\mathbf{GVS} + \mathbf{XF_{VX}US},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}] - r[\mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}],
\end{aligned} \tag{3.15}$$

from (1.18), (3.3), and (3.9) that

$$\begin{aligned}
&\max_{\mathbf{U}} r[\mathbf{GVS} + \mathbf{XF_{VX}US},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}] \\
&\quad = \max_{\mathbf{U}} r([\mathbf{GVS},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}] + \mathbf{XF_{VX}U}[\mathbf{S},\ \mathbf{0},\ \mathbf{0}]) \\
&\quad = \min\{r[\mathbf{GVS},\ \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2},\ \mathbf{XF_{VX}}],\quad r(\mathbf{S}) + r[\mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}]\} \\
&\quad = \min\{r[\mathbf{GV},\ \mathbf{XF_{VX}}],\ r(\mathbf{S}) + r[\mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}]\},
\end{aligned} \tag{3.16}$$

and from (1.13) that

$$r[\mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}] = r(\mathbf{N}) - r(\mathbf{VX}_1) - r(\mathbf{VX}_2). \tag{3.17}$$

Combining (3.15) and (3.16) gives

$$\begin{aligned}
&\max_{\mathbf{U}} \min_{\mathbf{U}_1,\mathbf{U}_2} r(\mathbf{GVS} + \mathbf{XF_{VX}US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S} - \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}\mathbf{U}_2\mathbf{S}) \\
&\quad = \min\{r[\mathbf{GV},\ \mathbf{XF_{VX}}] - r[\mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}],\ r(\mathbf{S})\} \\
&\quad = r[\mathbf{GV},\ \mathbf{XF_{VX}}] - r[\mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1},\ \mathbf{X}_2\mathbf{F}_{\mathbf{VX}_2}] \\
&\quad = 2r(\mathbf{X}_1'\mathbf{VX}_2) + r(\mathbf{X}) - r(\mathbf{N}) \quad \text{(by (3.11), (3.13), and (3.17))}.
\end{aligned}$$

Hence, (3.14) is equivalent to $2r(\mathbf{X}_1'\mathbf{VX}_2) + r(\mathbf{X}) = r(\mathbf{N})$. □

Proof of Theorem 2.4. It can be seen from (1.8) that $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ is unique if and only if $\mathbf{XF_{VX}} = \mathbf{0}$, which is equivalent to $r(\mathbf{X}) = r(\mathbf{VX})$ by (3.5). In this case, $E[\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}(\mathbf{X}'\mathbf{VX})^{+}\mathbf{X}'\mathbf{VX}\boldsymbol{\beta} = \mathbf{X}\boldsymbol{\beta}$, as required for (a). The results in (b) can be shown similarly.

Proof of Theorem 2.5. The range equality $\mathscr{R}(\mathbf{X}'\mathbf{V}) = \mathscr{R}(\mathbf{X}')$ can be rewritten in the partitioned form $\mathscr{R}\begin{bmatrix}\mathbf{X}_1'\mathbf{V}\\ \mathbf{X}_2'\mathbf{V}\end{bmatrix} = \mathscr{R}\begin{bmatrix}\mathbf{X}_1'\\ \mathbf{X}_2'\end{bmatrix}$, which implies both $\mathscr{R}(\mathbf{X}_1'\mathbf{V}) = \mathscr{R}(\mathbf{X}_1')$ and $\mathscr{R}(\mathbf{X}_2'\mathbf{V}) = \mathscr{R}(\mathbf{X}_2')$. Hence, the uniqueness of $\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})$ implies the uniqueness of both $\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)$ and $\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)$, as required for (a). The result in (b) is derived from (2.4). It is easy to find from (3.12) that

$$r[\mathbf{X}_1(\mathbf{X}_1'\mathbf{VX}_1)^{+}\mathbf{X}_1'\mathbf{V\Sigma VX}_2(\mathbf{X}_2'\mathbf{VX}_2)^{+}\mathbf{X}_2')] = r(\mathbf{X}_1'\mathbf{V\Sigma VX}_2).$$

The result in (c) is a simple consequence of this rank equality. The result in (d) follows from (a) and Theorem 2.1. □

Proof of Corollary 2.1. The equivalence of (a) and (e) follows from Theorem 2.1. The equivalence of (b) and (e) follows from (2.7) and (2.8). The equivalence of (c) and (e) follows from (2.10). It is also easy to find from (3.13) that

$$\begin{aligned} & r(Cov[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] - Cov[\mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1)] - Cov[\mathrm{WLSE}_{\mathscr{M}_2}(\mathbf{X}_2\boldsymbol{\beta}_2)]) \\ &= r[\mathbf{X}(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}' - \mathbf{X}_1(\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1' - \mathbf{X}_2(\mathbf{X}_2'\Sigma^{-1}\mathbf{X}_2)^{-1}\mathbf{X}_2'] \\ &= 2r(\mathbf{X}_1'\Sigma^{-1}\mathbf{X}_2). \end{aligned}$$

Hence, (d) and (e) are equivalent. □

Proof of Theorem 2.6. Note from (1.6) and (1.10) that

$$\mathrm{WLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) - \mathrm{WLSE}_{\mathscr{M}_1}(\mathbf{X}_1\boldsymbol{\beta}_1) = (\mathbf{GV} + \mathbf{XF_{VX}U} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1)\mathbf{y},$$

where $\mathbf{G} = \mathbf{X}(\mathbf{X}'\mathbf{VX})^{+}\mathbf{X}' - \mathbf{X}_1(\mathbf{X}_1'\mathbf{VX}_1)^{+}\mathbf{X}_1'$, and $\mathbf{U}$ and $\mathbf{U}_1$ are arbitrary. Hence it can be seen from (2.2) that (2.11) holds if and only if there exist matrices $\mathbf{U}$ and $\mathbf{U}_1$ such that

$$(\mathbf{GV} + \mathbf{XF_{VX}U} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1)\mathbf{y} = \mathbf{0} \quad \text{for all } \mathbf{y} \in \mathscr{R}[\mathbf{X}, \Sigma],$$

that is, there exist $\mathbf{U}$ and $\mathbf{U}_1$ such that

$$\left(\mathbf{GV} + [\mathbf{XF_{VX}}, \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}]\begin{bmatrix}\mathbf{U}\\ -\mathbf{U}_1\end{bmatrix}\right)\mathbf{S} = \mathbf{0}, \tag{3.18}$$

where $\mathbf{S} = [\mathbf{X}, \boldsymbol{\Sigma}]$. Rewrite (3.18) as

$$\mathbf{AZS} = -\mathbf{GVS}, \tag{3.19}$$

where $\mathbf{A} = [\mathbf{XF_{VX}}, \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}]$ and $\mathbf{Z} = [\mathbf{U}', -\mathbf{U}_1']'$. It can be derived from (1.20) that the equation in (3.19) is solvable for $\mathbf{Z}$ if and only if

$$r[\mathbf{GVS}, \mathbf{A}] = r(\mathbf{A}). \tag{3.20}$$

It is also easy to verify $\mathscr{R}(\mathbf{G}) = \mathscr{R}(\mathbf{G}') \subseteq \mathscr{R}(\mathbf{S})$, so that $\mathscr{R}(\mathbf{GVS}) = \mathscr{R}(\mathbf{GV})$ holds. In this case, we can derive from (1.16), (1.20), and (3.3) that (3.20) is equivalent to

$$r[\mathbf{GV}, \mathbf{XF_{VX}}] = r(\mathbf{XF_{VX}}) = r(\mathbf{X}) - r(\mathbf{VX}). \tag{3.21}$$

By (1.13) and elementary block matrix operations,

$$r[\mathbf{GV}, \mathbf{XF_{VX}}] = r\begin{bmatrix}\mathbf{GV} & \mathbf{X}\\ \mathbf{0} & \mathbf{VX}\end{bmatrix} - r(\mathbf{VX}) = r(\mathbf{VGV}) + r(\mathbf{X}) - r(\mathbf{VX}). \tag{3.22}$$

Thus, (3.21) is equivalent to $\mathbf{VGV} = \mathbf{0}$, which is, in turn, is equivalent to (c). Under the conditions in (3.12), applying (1.17) to $\mathbf{VGV}$ gives

$$r(\mathbf{VGV}) = r(\mathbf{Z} - \mathbf{Z}_1) = r(\mathbf{Z}) - r(\mathbf{Z}_1) = r(\mathbf{VX}) - r(\mathbf{VX}_1). \tag{3.23}$$

Thus, (3.21) is equivalent to $r(\mathbf{VX}) = r(\mathbf{VX}_1)$, which is also equivalent to $\mathscr{R}(\mathbf{VX}_2) \subseteq \mathscr{R}(\mathbf{VX}_1)$ by (1.15). Hence, (a) and (e) of the theorem are equivalent.

It can be seen from (3.18) that the set inclusion in (b) holds if and only if there exists a $\mathbf{U}$ such that

$$\min_{\mathbf{U}} r(\mathbf{GVS} + \mathbf{XF}_{\mathbf{VX}}\mathbf{US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S}) = 0 \tag{3.24}$$

for any $\mathbf{U}_1$. By (1.19), (3.3), (3.5), (3.22), and (3.23),

$$\begin{aligned}
&\min_{\mathbf{U}} r(\mathbf{GVS} + \mathbf{XF}_{\mathbf{VX}}\mathbf{US} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S}) \\
&\quad = r[\mathbf{GVS} - \mathbf{X}_1\mathbf{F}_{\mathbf{VX}_1}\mathbf{U}_1\mathbf{S}, \mathbf{XF}_{\mathbf{VX}}] - r(\mathbf{XF}_{\mathbf{VX}}) \\
&\quad = r[\mathbf{GV}, \mathbf{XF}_{\mathbf{VX}}] - r(\mathbf{XF}_{\mathbf{VX}}) \\
&\quad = r(\mathbf{VX}) - r(\mathbf{VX}_1).
\end{aligned}$$

Thus (3.24) is equivalent to (d). □

Acknowledgments

The authors are grateful to an anonymous referee for his/her helpful comments and constructive suggestions on an earlier version of the article. This research has been supported by the Shanghai Pujiang Program and by a research grant (#10630) from Natural Sciences and Engineering Research Council of Canada.

References

Bhimasankaram, P., Saharay, R. (1997). On a partitioned linear model and some associated reduced models. *Linear Algebra Appl.* 264:329–339.

Chu, K. L., Isotalo, J., Puntanen, S., Styan, G. P. H. (2004). On decomposing the Watson efficiency of ordinary least squares in a partitioned weakly singular linear model. *Sankhyā, Ser. A* 66:634–651.

Groß, J., Puntanen, S. (2000). Estimations under a general partitioned linear model. *Linear Algebra Appl.* 321:131–144.

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

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

Nurhonen, M., Puntanen, S. (1992). A property of partitioned generalized regression. *Commun. Statist. Theor. Meth.* 21:1579–1583.

Puntanen, S., Styan, G. P. H., Tian, Y. (2005). Three rank formulas associated with the covariance matrices of the BLUE and the OLSE in the general linear model. *Econometric Theor.* 21:659–664.

Qian, H., Tian, Y. (2006). Partially superfluous observations. *Econometric Theor.* 22:529–536.

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

Rao, C. R. (1973). Representations of best linear unbiased estimators in the Gauss–Markoff model with a singular dispersion matrix. *J. Multivariate Anal.* 3:276–292.

Rao, C. R., Mitra, S. K. (1971a). Further contributions to the theory of generalized inverse of matrices and its applications. *Sankhyā, Ser. A* 33:289–300.

Rao, C. R., Mitra, S. K. (1971b). *Generalized Inverse of Matrices and Its Applications*. New York: Wiley.

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.

Tian, Y., Wiens, D. P. (2006). On equality and proportionality of ordinary least-squares, weighted least-squares and best linear unbiased estimators in the general linear model. *Statist. Probab. Lett.* 76:1265–1272.

Werner, H. J., Yapar, C. (1995). More on partitioned possibly restricted linear regression. *Multivariate Statistics and Matrices in Statistics. New Trends in Probability and Statistics*. Vol. 3, Proceedings of the 5th Tartu Conference, Tartu, pp. 57–66.

Werner, H. J., Yapar, C. (1996). A BLUE decomposition in the general linear regression model. *Linear Algebra Appl.* 237/238:395–404.

Zhang, B., Liu, B., Lu, C. (2004). A study of the equivalence of the BLUEs between a partitioned singular linear model and its reduced singular linear models. *Acta Math. Sinica, Ser. B* 20:557–568.