DECOMPOSITION APPROACHES OF A CONSTRAINED GENERAL LINEAR MODEL WITH FIXED PARAMETERS[*]

BO JIANG[†] AND YONGGE TIAN[‡]

Abstract. The well-known ordinary least-squares estimators (OLSEs) and the best linear unbiased estimators (BLUEs) under linear regression models can be represented by certain closed-form formulas composed by the given matrices and their generalized inverses in the models. This paper provides a general algebraic approach to relationships between OLSEs and BLUEs of the whole and partial mean parameter vectors in a constrained general linear model (CGLM) with fixed parameters by using a variety of matrix analysis tools on generalized inverses of matrices and matrix rank formulas. In particular, it establishes a variety of necessary and sufficient conditions for OLSEs to be BLUEs under a CGLM, which include many reasonable statistical interpretations on the equalities of OLSEs and BLUEs of parameter space in the CGLM. The whole work shows how to effectively establish matrix equalities composed by matrices and their generalized inverses and how to use them when characterizing performances of estimators of parameter spaces in linear models under most general assumptions.

Key words. General linear model, Restriction, Estimability, OLSE, BLUE, Equality, Statistical interpretation.

AMS subject classifications. 15A03, 15A09, 62H12, 62F30.

1. Introduction. We consider the general linear model (GLM) defined by

$$\mathscr{M}: \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{X}_k\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}, \quad \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad \mathrm{D}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{\Sigma}, \tag{1.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observable random variables, $\mathbf{X} = [\,\mathbf{X}_1, \ldots, \mathbf{X}_k\,]$ is an $n \times p$ model matrix of arbitrary rank, $\mathbf{X}_1, \ldots, \mathbf{X}_k$ are $n \times p_1, \ldots, n \times p_k$ known matrices of arbitrary ranks with $p = p_1 + \cdots + p_k$, $\boldsymbol{\beta} = [\,\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_k'\,]'$ is a $p \times 1$ vector of fixed but unknown parameters and $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_k$ are $p_1 \times 1, \ldots, p_k \times 1$ vectors, $\boldsymbol{\varepsilon}$ is an $n \times 1$ vector of randomly distributed error terms with the expectation $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and the dispersion matrix $\mathrm{D}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{\Sigma}$, in which $\boldsymbol{\Sigma}$ is an $n \times n$ known nonnegative definite matrix of arbitrary rank, and σ^2 is an arbitrary positive scaling factor. One of the main objectives in the statistical inference of a GLM is to establish various estimators of the parameter space in the GLM and to characterize mathematical and statistical properties and features of these estimators under various assumptions. During this approach, statisticians are often interested in the connections of different estimators, and especially, are interested in establishing possible equalities between estimators.

Assume further that the unknown parameter vector $\boldsymbol{\beta}$ in (1.1) satisfies a consistent linear matrix equation $\mathbf{A}\boldsymbol{\beta} = \mathbf{A}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{A}_k\boldsymbol{\beta}_k = \mathbf{b}$, where $\mathbf{A} = [\,\mathbf{A}_1, \ldots, \mathbf{A}_k\,]$, $\mathbf{A}_1, \ldots, \mathbf{A}_k$, and $\mathbf{b}$ are $m \times p$, $m \times p_1, \ldots, m \times p_k$, and $m \times 1$ matrices, respectively. Then, we obtain a constrained general linear model (CGLM)

$$\widetilde{\mathscr{M}}: \quad \begin{cases} \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \mathbf{X}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{X}_k\boldsymbol{\beta}_k + \boldsymbol{\varepsilon}, \\ \mathbf{A}\boldsymbol{\beta} = \mathbf{A}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{A}_k\boldsymbol{\beta}_k = \mathbf{b}, \quad \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad \mathrm{D}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{\Sigma}. \end{cases} \tag{1.2}$$

[†]College of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai, China (jiangboliumengyu@gmail.com).

[‡]CEMA, Central University of Finance and Economics, Beijing, China (yongge.tian@gmail.com).

The linear matrix equation in (1.2) is often available as extraneous information for the unknown parameter vector $\boldsymbol{\beta}$ to satisfy, which is therefore an integral part of the CGLM about the parameter vector $\boldsymbol{\beta}$ and thus should ideally be utilized in any estimation procedure of the parameter space in (1.1). In regression analysis, a linear regression model is often written as sums of partial regressors in order to identify the most important regressors. Through the partitions of regressors in a regression model, it is quite convenient to determine the roles of partial regressors, and to derive estimators of partial parameter spaces in such a CGLM.

Eq. (1.2) is a standard form of linear statistical models with linear parameter restrictions. This kind of CGLMs are a classic object of study in regression theory and occur in most textbooks on linear regression models. In statistical analysis of CGLMs, parameter constraints are usually handled by transforming the explicitly constrained model into an implicitly constrained model that takes the form of a reduced-parameter, unconstrained model; see, e.g., [8, 9, 24, 37, 46, 48, 59, 61]. The commonly-used treatment of (1.2) is merging the main model and the linear restriction equation as the following implicitly-constrained model

$$\widetilde{\mathscr{M}}: \quad \widetilde{\mathbf{y}} = \widetilde{\mathbf{X}}\boldsymbol{\beta} + \widetilde{\boldsymbol{\varepsilon}} = \widetilde{\mathbf{X}}_1\boldsymbol{\beta}_1 + \cdots + \widetilde{\mathbf{X}}_k\boldsymbol{\beta}_k + \widetilde{\boldsymbol{\varepsilon}}, \quad \mathrm{E}(\widetilde{\boldsymbol{\varepsilon}}) = \mathbf{0}, \quad \mathrm{D}(\widetilde{\boldsymbol{\varepsilon}}) = \sigma^2\widetilde{\boldsymbol{\Sigma}}, \tag{1.3}$$

where

$$\widetilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y} \\ \mathbf{b} \end{bmatrix}, \quad \widetilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix}, \quad \widetilde{\mathbf{X}}_i = \begin{bmatrix} \mathbf{X}_i \\ \mathbf{A}_i \end{bmatrix}, \quad \widetilde{\boldsymbol{\varepsilon}} = \begin{bmatrix} \boldsymbol{\varepsilon} \\ \mathbf{0} \end{bmatrix}, \quad \widetilde{\boldsymbol{\Sigma}} = \begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad i = 1, \ldots, k.$$

In the statistical inference of CGLMs, many estimators of parameter spaces can be defined from different optimality criteria. So that people are often interested in the relationship between different estimators, and especially, are interested in establishing various possible (decomposition) equalities between estimators. In this paper, we reconsider a general problem of establishing connections between the two well-known ordinary least-squares estimator (OLSE) and the best linear unbiased estimator (BLUE) of the full and partial parameter vectors in (1.2). Because OLSEs and BLUEs of parameter spaces in CGLMs are defined from two different optimality criteria, and thus, they have different performances in statistical inference. It is well known that OLSEs and BLUEs can be represented in analytical formulas consisting of the given matrices and their generalized inverses in linear models, and thus algebraic properties and performances of OLSEs and BLUEs can easily be derived from the analytical formulas. It has been noticed that OLSEs and BLUEs of the same parameter space in a linear regression model have some essential links, in particular, OLSEs and BLUEs are equivalent under some conditions, and the equivalences of OLSEs and BLUEs have many mathematical and statistical interpretations. The problem of establishing/characterizing relations between OLSEs and BLUEs in linear regression theory was initialized and approached in the late 1940s from theoretical and applied points of view by many authors. For more and detailed information on this topic please refer to [3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 15, 19, 20, 21, 22, 23, 25, 26, 28, 29, 30, 31, 33, 35, 38, 41, 42, 43, 56, 57, 58, 60, 63, 66] and the references therein.

One of the fundamental problems in the theory of random vectors is to establish possible equalities between two random vectors from mathematical and statistical aspects. Assume that $\mathbf{L}_1\mathbf{y} + \mathbf{c}_1$ and $\mathbf{L}_2\mathbf{y} + \mathbf{c}_2$ are two linear statistics of the random vector $\mathbf{y}$ in (1.2). In order to characterize equalities between $\mathbf{L}_1\mathbf{y} + \mathbf{c}_1$ and $\mathbf{L}_2\mathbf{y} + \mathbf{c}_2$, we use the following three well-known criteria, which were intuitively applied in the statistical literature; see, e.g., [4, 10, 36, 41, 57, 58, 63, 64, 65].

DEFINITION 1.1. Let $\mathbf{y}$ be a random vector.

(i) The equality $\mathbf{L}_1\mathbf{y} + \mathbf{c}_1 = \mathbf{L}_2\mathbf{y} + \mathbf{c}_2$ is said to hold definitely if $\mathbf{L}_1 = \mathbf{L}_2$ and $\mathbf{c}_1 = \mathbf{c}_2$.

(ii) The equality $\mathbf{L}_1\mathbf{y}+\mathbf{c}_1 = \mathbf{L}_2\mathbf{y}+\mathbf{c}_2$ is said to hold with probability 1, i.e., both $\mathrm{E}(\mathbf{L}_1\mathbf{y}+\mathbf{c}_1-\mathbf{L}_2\mathbf{y}-\mathbf{c}_2) = \mathbf{0}$ and $\mathrm{D}(\mathbf{L}_1\mathbf{y} + \mathbf{c}_1 - \mathbf{L}_2\mathbf{y} - \mathbf{c}_2) = \mathbf{0}$ hold.

(iii) The expectation vectors and dispersion matrices of $\mathbf{L}_1\mathbf{y} + \mathbf{c}_1$ and $\mathbf{L}_2\mathbf{y} + \mathbf{c}_2$ are said to be equal, respectively, if both $\mathrm{E}(\mathbf{L}_1\mathbf{y} + \mathbf{c}_1) = \mathrm{E}(\mathbf{L}_2\mathbf{y} + \mathbf{c}_2)$ and $\mathrm{D}(\mathbf{L}_1\mathbf{y} + \mathbf{c}_1) = \mathrm{D}(\mathbf{L}_2\mathbf{y} + \mathbf{c}_2)$ hold.

These three types of equality are not necessarily equivalent since they are defined from different criteria. These, however, show that equalities of linear statistics under (1.2) can all be characterized by the corresponding linear and quadratic matrix equations. Because of the non-commutativity of matrix algebra, it is usually a challenging task to characterize various equalities of predictors/estimators in the statistical analysis of linear regression models. However, we are able to derive satisfactory conclusions, as demonstrated in Section 5, for many equalities between OLSEs and BLUEs under CGLMs.

Closed-form formulas of calculating OLSEs and BLUEs under GLMs can be established by some routine matrix operations, which can be found in most books and articles on GLMs. In particular, it is easy to verify from analytical expressions of OLSEs and BLUEs that the following two additive decompositions

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}_1\boldsymbol{\beta}_1) + \cdots + \mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}_k\boldsymbol{\beta}_k), \tag{1.4}$$

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}_1\boldsymbol{\beta}_1) + \cdots + \mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}_k\boldsymbol{\beta}_k) \tag{1.5}$$

hold naturally under the conditions that the $\mathbf{X}_1\boldsymbol{\beta}_1, \ldots, \mathbf{X}_k\boldsymbol{\beta}_k$ are estimable under (1.1), respectively. On the other hand, the equalities of OLSEs and BLUEs have many different statistical interpretations, and are the criteria of comparing optimality of estimators in statistical inference. Under this consideration, it would be of interest to compare both sides of the two additive decomposition equalities in (1.4) and (1.5) and to establish possible links of OLSEs and BLUEs of the full and partial mean parameters. In this paper, we approach the following two problems of establishing equalities between the OLSEs and BLUEs of the whole and partial mean parameter vectors in (1.3):

(I) establishing necessary and sufficient conditions for the following equalities

$$\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \quad \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)$$

to hold definitely (with probability 1), $i = 1, \ldots, k$;

(II) showing that the following four statistical assertions
   (a) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})$ holds definitely (with probability 1),
   (b) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})$ holds definitely (with probability 1),
   (c) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)$ holds definitely (with probability 1), $i = 1, \ldots, k$,
   (d) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)$ holds definitely (with probability1), $i = 1, \ldots, k$
   are equivalent.

This kind of estimator equalities have many different valuable statistical interpretations, and are not rare to see in the statistical inference of CGLMs. Because OLSEs and BLUEs of parameter spaces in CGLMs can be calculated from given matrices and vectors in the models and are often represented by certain closed-form formulas composed by given matrices and vectors in the CGLMs, the approaches we take to the above problems is in fact to establish and characterize certain matrix equalities that involve the given matrices and their generalized inverses in (1.2). Thus, we can use various matrix analysis tools to characterize the equalities in (I) and (II). Approaches to the equivalences of OLSEs and BLUEs of full and partial mean parameter vectors in GLMs were initiated in [64]; connections among OLSEs and BLUEs of whole and partial parameters under multiple partitioned linear models were considered in [65]; equivalences of OLSEs

and BLUEs of full and partial mean parameter vectors in CGLMs with separated parameter restrictions were approached in [62]; whereas the above two problems for $i = 2$ were investigated in [27].

The following terminology and notation are used throughout the paper. The symbol $\mathbb{R}^{m \times n}$ stands for the collection of all $m \times n$ real matrices. The symbols $\mathbf{A}'$, $r(\mathbf{A})$, and $\mathscr{R}(\mathbf{A})$ stand for the transpose, the rank, and the range (column space) of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, respectively; $\mathbf{I}_m$ stands for the identity matrix of order m. Two symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ of the same size are said to satisfy the inequality $\mathbf{A} \succcurlyeq \mathbf{B}$ in the Löwner partial ordering if $\mathbf{A} - \mathbf{B}$ is nonnegative definite. The Moore–Penrose generalized inverse of $\mathbf{A} \in \mathbb{R}^{m \times n}$, denoted by $\mathbf{A}^+$, is defined to be the unique solution $\mathbf{G}$ satisfying the four matrix equations $\mathbf{AGA} = \mathbf{A}$, $\mathbf{GAG} = \mathbf{G}$, $(\mathbf{AG})' = \mathbf{AG}$, and $(\mathbf{GA})' = \mathbf{GA}$. Further, denote $\mathbf{P_A} = \mathbf{AA}^+$, $\mathbf{A}^\perp = \mathbf{E_A} = \mathbf{I}_m - \mathbf{AA}^+$, and $\mathbf{F_A} = \mathbf{E}_{\mathbf{A}'} = \mathbf{I}_n - \mathbf{A}^+\mathbf{A}$.

2. A useful inclusion on matrix rank and its importance. Much important statistical theory and many statistical methods make use of linear algebra and matrix theory, and many statisticians have contributed to the advancement of matrix theory from its very early days. In particular, formulas and algebraic techniques for handling matrices in linear algebra and matrix theory play important roles in the derivations of estimators and characterizations of performances of the estimators and predictors in statistical analysis of linear regression models. As remarked in [49], a good starting point for the entry of matrices into statistical sciences was in 1930s. It is now a routine procedure to use given vectors, matrices, and their inverses/generalized inverses to formulate various estimators of parameter spaces and to make the corresponding statistical inference.

Let's recall that the rank of matrix is a core concept in linear algebra, which is the most significant finite nonnegative integer in reflecting intrinsic properties of matrix, and thus is a cornerstone in matrix theory. The mathematical prerequisites for understanding ranks of matrices are minimal and do not go beyond elementary linear algebra, while many simple and classic formulas for calculating ranks of matrices can be found in most textbooks of linear algebra. The whole work in this paper is based on the effective use of the matrix rank formulas when simplifying various matrix expressions and establishing matrix equalities occurred in statistical inference of CGLMs. In order to simplify various matrix equalities composed by the Moore–Penrose inverses of matrices, we need to use the matrix rank formulas and the equivalent facts in the following two lemmas.

LEMMA 2.1 ([32]). *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$, $\mathbf{C} \in \mathbb{R}^{l \times n}$, *and* $\mathbf{D} \in \mathbb{R}^{l \times k}$. *Then*

$$r[\, \mathbf{A},\, \mathbf{B}\,] = r(\mathbf{A}) + r(\mathbf{E_A B}) = r(\mathbf{B}) + r(\mathbf{E_B A}), \tag{2.6}$$

$$r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) + r(\mathbf{CF_A}) = r(\mathbf{C}) + r(\mathbf{AF_C}), \tag{2.7}$$

$$r\begin{bmatrix} \mathbf{AA}' & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{bmatrix} = r[\, \mathbf{A},\, \mathbf{B}\,] + r(\mathbf{B}). \tag{2.8}$$

If $\mathscr{R}(\mathbf{B}) \subseteq \mathscr{R}(\mathbf{A})$ *and* $\mathscr{R}(\mathbf{C}') \subseteq \mathscr{R}(\mathbf{A}')$, *then*

$$r\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = r(\mathbf{A}) + r(\mathbf{D} - \mathbf{CA}^+\mathbf{B}). \tag{2.9}$$

In addition, the following results hold.

(i) $r[\, \mathbf{A},\, \mathbf{B}\,] = r(\mathbf{A}) \Leftrightarrow \mathscr{R}(\mathbf{B}) \subseteq \mathscr{R}(\mathbf{A}) \Leftrightarrow \mathbf{AA}^+\mathbf{B} = \mathbf{B} \Leftrightarrow \mathbf{E_A B} = \mathbf{0}$.

(ii) $r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) \Leftrightarrow \mathscr{R}(\mathbf{C}') \subseteq \mathscr{R}(\mathbf{A}') \Leftrightarrow \mathbf{C}\mathbf{A}^+\mathbf{A} = \mathbf{C} \Leftrightarrow \mathbf{C}\mathbf{F_A} = \mathbf{0}$.

(iii) $r[\,\mathbf{A},\ \mathbf{B}\,] = r(\mathbf{A}) + r(\mathbf{B}) \Leftrightarrow \mathscr{R}(\mathbf{A}) \cap \mathscr{R}(\mathbf{B}) = \{\mathbf{0}\} \Leftrightarrow \mathscr{R}[(\mathbf{E_A}\mathbf{B})'] = \mathscr{R}(\mathbf{B}') \Leftrightarrow \mathscr{R}[(\mathbf{E_B}\mathbf{A})'] = \mathscr{R}(\mathbf{A}')$.

(iv) $r\begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) + r(\mathbf{C}) \Leftrightarrow \mathscr{R}(\mathbf{A}') \cap \mathscr{R}(\mathbf{C}') = \{\mathbf{0}\} \Leftrightarrow \mathscr{R}(\mathbf{C}\mathbf{F_A}) = \mathscr{R}(\mathbf{C}) \Leftrightarrow \mathscr{R}(\mathbf{A}\mathbf{F_C}) = \mathscr{R}(\mathbf{A})$.

(v) $r(\,\mathbf{A} + \mathbf{B}\,) = r(\mathbf{A}) + r(\mathbf{B}) \Leftrightarrow \mathscr{R}(\mathbf{A}) \cap \mathscr{R}(\mathbf{B}) = \{\mathbf{0}\}$ and $\mathscr{R}(\mathbf{A}') \cap \mathscr{R}(\mathbf{B}') = \{\mathbf{0}\}$ under $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$.

LEMMA 2.2 ([55]). *Suppose that* $\mathscr{R}(\mathbf{A}) \subseteq \mathscr{R}(\mathbf{B}_1)$, $\mathscr{R}(\mathbf{C}_2) \subseteq \mathscr{R}(\mathbf{C}_1)$, $\mathscr{R}(\mathbf{A}') \subseteq \mathscr{R}(\mathbf{C}_1')$, and $\mathscr{R}(\mathbf{B}_2') \subseteq \mathscr{R}(\mathbf{B}_1')$. Then

$$r(\mathbf{B}_2\mathbf{B}_1^+\mathbf{A}\mathbf{C}_1^+\mathbf{C}_2) = r\begin{bmatrix} \mathbf{A} & \mathbf{B}_1 & \mathbf{0} \\ \mathbf{C}_1 & \mathbf{0} & \mathbf{C}_2 \\ \mathbf{0} & \mathbf{B}_2 & \mathbf{0} \end{bmatrix} - r(\mathbf{B}_1) - r(\mathbf{C}_1). \tag{2.10}$$

LEMMA 2.3 ([40]). *The linear matrix equation* $\mathbf{A}\mathbf{X} = \mathbf{B}$ *is consistent if and only if* $r[\,\mathbf{A},\ \mathbf{B}\,] = r(\mathbf{A})$, *or equivalently,* $\mathbf{A}\mathbf{A}^+\mathbf{B} = \mathbf{B}$. *In this case, the general solution of the equation can be written as* $\mathbf{X} = \mathbf{A}^+\mathbf{B} + (\mathbf{I} - \mathbf{A}^+\mathbf{A})\mathbf{U}$, *where* $\mathbf{U}$ *is an arbitrary matrix.*

With the supports of the formulas in Lemmas 2.1–2.3, we are able to covert Problems (I) and (II) in Section 1 into certain algebraic problems of characterizing matrix equalities composed by the given matrices in the models and their generalized inverses, and to derive analytical solutions of the problems by using the methods of matrix equations, matrix rank formulas, and various tricky partitioned matrix calculations.

3. Consistency and estimability of parameter spaces in CGLMs. Without loss of generality, we take $\sigma^2 = 1$ in (1.1)–(1.3) for the convenience of presentation below, because it does't involve in the main results in this paper. In what follows, we assume that $\widetilde{\mathscr{M}}$ in (1.2) is consistent, namely,

$$\widetilde{\mathbf{y}} \in \mathscr{R}[\,\widetilde{\mathbf{X}},\ \widetilde{\mathbf{\Sigma}}\,] \ \text{ holds with probability 1,}$$

see [44, 45] for its expositions. In this case, (1.1) is consistent as well, that is, $\mathbf{y} \in \mathscr{R}[\,\mathbf{X},\ \mathbf{\Sigma}\,]$ holds with probability 1. Let

$$\mathbf{S}_i = [\,\mathbf{0}, \ldots, \mathbf{K}_i, \ldots, \mathbf{0}\,], \quad \mathbf{T}_i = [\,\mathbf{K}_1, \ldots, \mathbf{K}_{i-1}, \mathbf{0}, \mathbf{K}_{i+1}, \ldots, \mathbf{K}_k\,],$$

$$\mathbf{Y}_i = [\,\mathbf{0}, \ldots, \mathbf{X}_i, \ldots, \mathbf{0}\,], \quad \mathbf{Z}_i = [\,\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}, \mathbf{0}, \mathbf{X}_{i+1}, \ldots, \mathbf{X}_k\,],$$

$$\mathbf{W}_i = [\,\mathbf{0}, \ldots, \widetilde{\mathbf{X}}_i, \ldots, \mathbf{0}\,], \quad \mathbf{V}_i = [\,\widetilde{\mathbf{X}}_1, \ldots, \widetilde{\mathbf{X}}_{i-1}, \mathbf{0}, \widetilde{\mathbf{X}}_{i+1}, \ldots, \widetilde{\mathbf{X}}_k\,]$$

for $i = 1, \ldots, k$. Then, the arbitrary matrix $\mathbf{K}$, the model matrix $\mathbf{X}$, and $\widetilde{\mathbf{X}}$ in (1.2) and (1.3) can be decomposed as

$$\mathbf{K} = \mathbf{S}_i + \mathbf{T}_i = \mathbf{S}_1 + \cdots + \mathbf{S}_k, \quad \mathbf{X} = \mathbf{Y}_i + \mathbf{Z}_i = \mathbf{Y}_1 + \cdots + \mathbf{Y}_k \quad \text{and} \quad \widetilde{\mathbf{X}} = \mathbf{W}_i + \mathbf{V}_i = \mathbf{W}_1 + \cdots + \mathbf{W}_k$$

for $i = 1, \ldots, k$. Correspondingly, the vectors $\mathbf{K}_i\boldsymbol{\beta}_i$, $\mathbf{X}_i\boldsymbol{\beta}_i$, and $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ can be rewritten as

$$\mathbf{K}_i\boldsymbol{\beta}_i = \mathbf{S}_i\boldsymbol{\beta}, \quad \mathbf{X}_i\boldsymbol{\beta}_i = \mathbf{Y}_i\boldsymbol{\beta}, \quad \widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i = \mathbf{W}_i\boldsymbol{\beta}, \quad i = 1, \ldots, k.$$

We next introduce the definitions of the estimability of the parameter spaces in (1.1) and (1.2).

DEFINITION 3.1. Let $\mathscr{M}$ and $\widetilde{\mathscr{M}}$ be as given in (1.1) and (1.2), respectively, and let $\mathbf{K} \in \mathbb{R}^{t \times p}$ be given.

(i) The vector $\mathbf{K}\boldsymbol{\beta}$ of parameters is said to be estimable under (1.1) if there exists an $\mathbf{L} \in \mathbb{R}^{t \times n}$ such that $\mathrm{E}(\mathbf{L}\mathbf{y}) = \mathbf{K}\boldsymbol{\beta}$ holds under (1.1).

(ii) $\mathbf{K}\boldsymbol{\beta}$ is said to be estimable under (1.2) if there exist $\mathbf{L} \in \mathbb{R}^{t \times n}$ and $\mathbf{c} \in \mathbb{R}^{t \times 1}$ such that $\mathrm{E}(\mathbf{Ly}+\mathbf{c}) = \mathbf{K}\boldsymbol{\beta}$ holds under (1.2).

The unbiasedness of linear statistics with respect to given parameter spaces in CGLMs is an important property, but usually there are many unbiased estimators for the same vector of parameters. Considerable literature exists on the estimability of parameter spaces under CGLMs; see, e.g., [2, 34, 47, 50, 51, 52, 53, 61] for its expositions. Under the assumptions in (1.1) and (1.2), a variety of known classic results on the estimability of the vector $\mathbf{K}\boldsymbol{\beta}$ of parametric functions and its special cases are collected in the following lemma.

LEMMA 3.2. *Let $\mathscr{M}$ and $\widetilde{\mathscr{M}}$ be as given in* (1.1) *and* (1.2), *respectively, and let $\mathbf{K} \in \mathbb{R}^{t \times p}$ be given. Then, the following results hold.*

(i) $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.1) $\Leftrightarrow r\begin{bmatrix} \mathbf{X} \\ \mathbf{K} \end{bmatrix} = r(\mathbf{X}) \Leftrightarrow \mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$.

(ii) $\mathbf{X}\boldsymbol{\beta}$ *is always estimable under* (1.1).

(iii) $\mathbf{K}_i\boldsymbol{\beta}_i = \mathbf{S}_i\boldsymbol{\beta}$ *is estimable under* (1.1) $\Leftrightarrow \mathscr{R}(\mathbf{S}_i') \subseteq \mathscr{R}(\mathbf{X}'), i = 1, \ldots, k$.

(iv) *If all $\mathbf{K}_i\boldsymbol{\beta}_i$ are estimable under* (1.1), *then $\mathbf{K}\boldsymbol{\beta} = \mathbf{K}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{K}_k\boldsymbol{\beta}_k$ is estimable under* (1.1) *as well.*

(v) $\mathbf{X}_i\boldsymbol{\beta}_i = \mathbf{Y}_i\boldsymbol{\beta}$ *is estimable under* (1.1) $\Leftrightarrow \mathscr{R}(\mathbf{Y}_i') \subseteq \mathscr{R}(\mathbf{X}') \Leftrightarrow \mathscr{R}(\mathbf{Y}_i) \cap \mathscr{R}(\mathbf{Z}_i) = \mathscr{R}(\mathbf{X}_i) \cap \mathscr{R}(\mathbf{Z}_i) = \{\mathbf{0}\} \Leftrightarrow r(\mathbf{X}) = r(\mathbf{Y}_i) + r(\mathbf{Z}_i) = r(\mathbf{X}_i) + r(\mathbf{Z}_i), i = 1, \ldots, k$.

(vi) $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i = \mathbf{W}_i\boldsymbol{\beta}$ *is estimable under* (1.1) $\Leftrightarrow \mathscr{R}(\mathbf{W}_i') \subseteq \mathscr{R}(\mathbf{X}') \Leftrightarrow r(\mathbf{X}) = r(\mathbf{W}_i)+r(\mathbf{Z}_i) = r(\widetilde{\mathbf{X}}_i)+r(\mathbf{Z}_i), i = 1, \ldots, k$.

(vii) *If all $\mathbf{X}_i\boldsymbol{\beta}_i$ are estimable under* (1.1) $\Leftrightarrow r(\mathbf{X}) = r(\mathbf{X}_1) + \cdots + r(\mathbf{X}_k)$.

(viii) $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.2) $\Leftrightarrow \mathscr{R}[(\mathbf{KF_A})'] \subseteq \mathscr{R}[(\mathbf{XF_A})'] \Leftrightarrow \mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\widetilde{\mathbf{X}}')$.

(ix) $\mathbf{X}\boldsymbol{\beta}$ *and $\widetilde{\mathbf{X}}\boldsymbol{\beta}$ are always estimable under* (1.2).

(x) $\mathbf{K}_i\boldsymbol{\beta}_i = \mathbf{S}_i\boldsymbol{\beta}$ *is estimable under* (1.2) $\Leftrightarrow \mathscr{R}(\mathbf{S}_i') \subseteq \mathscr{R}(\widetilde{\mathbf{X}}'), i = 1, \ldots, k$.

(xi) *If all $\mathbf{K}_i\boldsymbol{\beta}_i$ are estimable under* (1.2), *then $\mathbf{K}\boldsymbol{\beta} = \mathbf{K}_1\boldsymbol{\beta}_1 + \cdots + \mathbf{K}_k\boldsymbol{\beta}_k$ is estimable under* (1.2) *as well.*

(xii) $\mathbf{X}_i\boldsymbol{\beta}_i = \mathbf{Y}_i\boldsymbol{\beta}$ *is estimable under* (1.2) $\Leftrightarrow \mathscr{R}(\mathbf{Y}_i') \subseteq \mathscr{R}(\widetilde{\mathbf{X}}'), i = 1, \ldots, k$.

(xiii) $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i = \mathbf{W}_i\boldsymbol{\beta}$ *is estimable under* (1.2) $\Leftrightarrow \mathscr{R}(\mathbf{W}_i') \subseteq \mathscr{R}(\widetilde{\mathbf{X}}') \Leftrightarrow \mathscr{R}(\mathbf{W}_i) \cap \mathscr{R}(\mathbf{V}_i) = \mathscr{R}(\widetilde{\mathbf{X}}_i) \cap \mathscr{R}(\mathbf{V}_i) = \{\mathbf{0}\} \Leftrightarrow r(\widetilde{\mathbf{X}}) = r(\mathbf{W}_i) + r(\mathbf{V}_i) = r(\widetilde{\mathbf{X}}_i) + r(\mathbf{V}_i), i = 1, \ldots, k$.

(xiv) *If all $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ are estimable under* (1.2) $\Leftrightarrow r(\widetilde{\mathbf{X}}) = r(\widetilde{\mathbf{X}}_1) + \cdots + r(\widetilde{\mathbf{X}}_k)$.

(xv) *If $\mathbf{K}_i\boldsymbol{\beta}_i$ is estimable under* (1.1), *then $\mathbf{K}_i\boldsymbol{\beta}_i$ is estimable under* (1.2) *as well, $i = 1, \ldots, k$.*

(xvi) *If $\mathbf{X}_i\boldsymbol{\beta}_i$ is estimable under* (1.1), *then $\mathbf{X}_i\boldsymbol{\beta}_i$ is estimable under* (1.2) *as well, $i = 1, \ldots, k$.*

(xvii) *If $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ is estimable under* (1.1), *then $\mathbf{X}_i\boldsymbol{\beta}_i$ is estimable under* (1.2) *as well, $i = 1, \ldots, k$.*

Proof. Result (i) is well known; see, e.g., [2]. Results (ii)–(iv) and (vi) follow from (i). Results (v) and (vii) follow from [65]. By the definition of estimability, $\mathbf{K}\boldsymbol{\beta}$ is estimable under (1.2) if there exists a linear statistic $\mathbf{Ly} + \mathbf{c}$, where $\mathbf{L} \in \mathbb{R}^{t \times n}$ and $\mathbf{c} \in \mathbb{R}^{t \times 1}$, such that $E(\mathbf{Ly} + \mathbf{c}) = \mathbf{LX}\boldsymbol{\beta} + \mathbf{c} = \mathbf{K}\boldsymbol{\beta}$ holds subject to $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$. From Lemma 2.3, the general solution of the consistent matrix equation $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$ is $\boldsymbol{\beta} = \mathbf{A}^+\mathbf{b} + \mathbf{F_A}\boldsymbol{\gamma}$, where $\boldsymbol{\gamma}$ is arbitrary. Hence, $\mathbf{LX}\boldsymbol{\beta} + \mathbf{c} = \mathbf{K}\boldsymbol{\beta}$ is equivalent to $\mathbf{LXA}^+\mathbf{b} + \mathbf{LXF_A}\boldsymbol{\gamma} + \mathbf{c} = \mathbf{KA}^+\mathbf{b} + \mathbf{KF_A}\boldsymbol{\gamma}$. This equality holds for all $\boldsymbol{\gamma}$ if and only if $\mathbf{LXA}^+\mathbf{b} + \mathbf{c} = \mathbf{KA}^+\mathbf{b}$ and $\mathbf{LXF_A} = \mathbf{KF_A}$. From Lemma 2.3, $\mathbf{LXF_A} = \mathbf{KF_A}$ is consistent if and only if the last two statements in (viii) hold. Results (ix)–(xii) follow from (viii). Results (xiii) and (xiv) are similar to (v) and (vii). Results (xv)–(xvii) follow from (iii), (v), (vi), (x), (xii), and (xiii). □

4. Closed-form formulas of OLSEs and BLUEs under CGLMs. The purpose of this section is to review the fundamentals of OLSEs and BLUEs in the context of linear regression analysis. The method of least squares in statistics is a standard technique used for estimating unknown parameters in CGLMs, which was first proposed as an algebraic procedure for solving overdetermined systems of equations by Gauss (in unpublished work) in 1795 and independently by Legendre in 1805, as remarked in [1, 17, 39, 54]. The notion of least-squares estimation is well established in the literature, and the definitions of the OLSEs and the BLUEs of parameter spaces in (1.1) and (1.2) are presented below.

DEFINITION 4.1. Let $\mathscr{M}$ and $\widetilde{\mathscr{M}}$ be as given in (1.1) and (1.2), respectively, and let $\mathbf{K} \in \mathbb{R}^{t \times p}$ be given.

(i) The OLSE of the parameter vector $\boldsymbol{\beta}$ in (1.1), denoted by $\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta})$, is defined to be

$$\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}) = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{4.11}$$

The OLSE of $\mathbf{K}\boldsymbol{\beta}$ under (1.1) is defined to be $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta})$.

(ii) The OLSE of the parameter vector $\boldsymbol{\beta}$ in (1.2), denoted by $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\boldsymbol{\beta})$, is defined to be

$$\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\boldsymbol{\beta}) = \underset{\mathbf{A}\boldsymbol{\beta}=\mathbf{b}}{\operatorname{argmin}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{4.12}$$

The OLSE of $\mathbf{K}\boldsymbol{\beta}$ under (1.2) is defined to be $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\boldsymbol{\beta})$.

Under the situation that there exist non-unique unbiased estimators for the same parameter space, it would seem naturally advantageous to seek such an unbiased estimator that has the smallest dispersion matrix among all the unbiased estimators. Thus, the unbiasedness and the smallest dispersion matrix of an estimator are two intrinsic requirements in statistical analysis and inference of linear regression analysis.

DEFINITION 4.2.   (i) If there exists an $\mathbf{L} \in \mathbb{R}^{t \times n}$ such that

$$\mathrm{E}(\mathbf{L}\mathbf{y} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0} \ \text{ and } \ \mathrm{D}(\mathbf{L}\mathbf{y} - \mathbf{K}\boldsymbol{\beta}) = \min \tag{4.13}$$

hold in the Löwner partial ordering, the linear statistic $\mathbf{L}\mathbf{y}$ is defined to be the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under (1.1), and is denoted by

$$\mathbf{L}\mathbf{y} = \mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}). \tag{4.14}$$

(ii) If there exists an $\mathbf{L} \in \mathbb{R}^{t \times n}$ and $\mathbf{c} \in \mathbb{R}^{t \times 1}$ such that

$$\mathrm{E}(\mathbf{L}\mathbf{y} + \mathbf{c} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0} \ \text{ and } \ \mathrm{D}(\mathbf{L}\mathbf{y} + \mathbf{c} - \mathbf{K}\boldsymbol{\beta}) = \min \tag{4.15}$$

hold in the Löwner partial ordering, the linear statistic $\mathbf{L}\mathbf{y} + \mathbf{c}$ is defined to be the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under (1.2), and is denoted by

$$\mathbf{L}\mathbf{y} + \mathbf{c} = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}). \tag{4.16}$$

It is well known that OLSEs and BLUEs are two widely-used estimators of parameter spaces in the statistical inference of CGLMs. These two types of estimator have a variety of simple and remarkable properties, and are regarded as orthodox representations of estimators under CGLMs. Notice from definitions that the OLSEs of $\mathbf{K}\boldsymbol{\beta}$ are independent of the dispersion matrix $\boldsymbol{\Sigma}$ in (1.1) and (1.2). Thus, it is of greatly interest from the theoretical and applied points of view to compare the OLSEs and the BLUEs of $\mathbf{K}\boldsymbol{\beta}$ under (1.1) and (1.2), as well as to give the efficiency for the OLSEs to be alternatives of the BLUEs of $\mathbf{K}\boldsymbol{\beta}$ under (1.1) and (1.2).

To account for general estimation problems of unknown parameters in CGLMs, it is common practice to first derive exact algebraic expressions of estimators of parameters spaces in the CGLMs. It is well known that the normal equation associated with (1.1) is $\mathbf{X}'\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'\mathbf{y}$; see, e.g., [18, p. 114] and [48, pp. 164–165]. From Lemma 2.3, the general solution of $\boldsymbol{\beta}$ of the equation and the corresponding $\mathbf{K}\boldsymbol{\beta}$ are given by

$$\boldsymbol{\beta} = \mathbf{X}^+\mathbf{y} + (\mathbf{I}_p - \mathbf{X}^+\mathbf{X})\mathbf{u}, \quad \mathbf{K}\boldsymbol{\beta} = \mathbf{K}\mathbf{X}^+\mathbf{y} + (\mathbf{K} - \mathbf{K}\mathbf{X}^+\mathbf{X})\mathbf{u}, \tag{4.17}$$

where $\mathbf{u}$ is arbitrary. In particular, the coefficient matrix of $\mathbf{u}$ satisfies $\mathbf{K} - \mathbf{K}\mathbf{X}^+\mathbf{X} = \mathbf{0}$ if and only if $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$ holds by Lemma 2.3. In this case, $\mathbf{K}\boldsymbol{\beta}$ can uniquely be written as $\mathbf{K}\mathbf{X}^+\mathbf{y}$. To sum up, we have the following well-known results.

LEMMA 4.3. *Let* $\mathbf{K} \in \mathbb{R}^{t \times p}$ *be given and assume that* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.1). *Then, the OLSE of* $\mathbf{K}\boldsymbol{\beta}$ *under* (1.1) *and its expectation and covariance matrix can be written as*

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathbf{X}^+\mathbf{y}, \tag{4.18}$$

$$\mathrm{E}[\,\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\boldsymbol{\beta}, \qquad \mathrm{D}[\,\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\mathbf{X}^+\boldsymbol{\Sigma}(\mathbf{K}\mathbf{X}^+)'. \tag{4.19}$$

In particular,

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P_X}\mathbf{y}, \ \ \mathrm{E}[\,\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\,] = \mathbf{X}\boldsymbol{\beta}, \ \ \mathrm{D}[\mathrm{OLSE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})\,] = \mathbf{P_X}\boldsymbol{\Sigma}\mathbf{P_X}. \tag{4.20}$$

Linear models with parameter constraints are usually handled by transforming into certain implicitly constrained model. The most popular transformations are based on model reduction and Lagrangian multipliers. From Lemma 2.3, the general solution to the matrix equation $\mathbf{A}\boldsymbol{\beta} = \mathbf{b}$ is $\boldsymbol{\beta} = \mathbf{A}^+\mathbf{b} + \mathbf{F_A}\boldsymbol{\gamma}$, where $\boldsymbol{\gamma}$ is an arbitrary vector. Then,

$$\mathbf{K}\boldsymbol{\beta} = \mathbf{K}\mathbf{A}^+\mathbf{b} + \widehat{\mathbf{K}}\boldsymbol{\gamma}, \quad \mathbf{X}\boldsymbol{\beta} = \mathbf{X}\mathbf{A}^+\mathbf{b} + \widehat{\mathbf{X}}\boldsymbol{\gamma}, \quad \widetilde{\mathbf{X}}\boldsymbol{\beta} = \widetilde{\mathbf{X}}\mathbf{A}^+\mathbf{b} + \widehat{\mathbf{Z}}\boldsymbol{\gamma}, \tag{4.21}$$

$$\mathbf{X}_i\boldsymbol{\beta}_i = \mathbf{Y}_i\boldsymbol{\beta} = \mathbf{Y}_i\mathbf{A}^+\mathbf{b} + \widehat{\mathbf{Y}}_i\boldsymbol{\gamma}, \quad \widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i = \mathbf{W}_i\boldsymbol{\beta} = \mathbf{W}_i\mathbf{A}^+\mathbf{b} + \widehat{\mathbf{W}}_i\boldsymbol{\gamma}, \tag{4.22}$$

where

$$\mathbf{Y}_i = [\,\mathbf{0}, \ldots, \mathbf{X}_i, \ldots, \mathbf{0}\,], \quad \mathbf{W}_i = [\,\mathbf{0}, \ldots, \widetilde{\mathbf{X}}_i, \ldots, \mathbf{0}\,], \tag{4.23}$$

$$\widehat{\mathbf{K}} = \mathbf{K}\mathbf{F_A}, \quad \widehat{\mathbf{X}} = \mathbf{X}\mathbf{F_A}, \quad \widehat{\mathbf{Z}} = \widetilde{\mathbf{X}}\mathbf{F_A}, \quad \widehat{\mathbf{Y}}_i = \mathbf{Y}_i\mathbf{F_A}, \quad \widehat{\mathbf{W}}_i = \mathbf{W}_i\mathbf{F_A} \tag{4.24}$$

for $i = 1, \ldots, k$. Substituting (4.21) into (1.2) yields the following reduced linear model

$$\mathscr{N} : \quad \mathbf{y} - \mathbf{X}\mathbf{A}^+\mathbf{b} = \widehat{\mathbf{X}}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \ \ \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \ \ \mathrm{D}(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma}. \tag{4.25}$$

Correspondingly,

$$\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathbf{A}^+\mathbf{b} + \mathrm{OLSE}_{\mathscr{N}}(\widehat{\mathbf{K}}\boldsymbol{\gamma}), \tag{4.26}$$

$$\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathbf{A}^+\mathbf{b} + \mathrm{BLUE}_{\mathscr{N}}(\widehat{\mathbf{K}}\boldsymbol{\gamma}). \tag{4.27}$$

The OLSE of $\widehat{\mathbf{K}}\boldsymbol{\gamma}$ under (4.25) can be derived from Lemma 4.3. Substituting it into (4.26) yields the OLSE of $\mathbf{K}\boldsymbol{\beta}$ under (1.2) as follows.

LEMMA 4.4. *Let* $\widetilde{\mathscr{M}}$ *and* $\mathscr{N}$ *be as given in* (1.2) *and* (4.25), *respectively, and let* $\mathbf{K} \in \mathbb{R}^{t \times p}$ *be given. Then, the following results hold.*

(i) *The following statements are equivalent*:

    (a) $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.2).

    (b) $\widehat{\mathbf{K}}\boldsymbol{\gamma}$ *is estimable under* (4.25).

    (c) $r\begin{bmatrix}\mathbf{K}\mathbf{F_A}\\\mathbf{X}\mathbf{F_A}\end{bmatrix} = r(\mathbf{X}\mathbf{F_A})$.

    (d) $\mathscr{R}[(\mathbf{K}\mathbf{F_A})'] \subseteq \mathscr{R}[(\mathbf{X}\mathbf{F_A})']$.

    (e) $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}[\,\mathbf{X}',\,\mathbf{A}'\,]$.

(ii) *Under* (i), *the OLSE of* $\mathbf{K}\boldsymbol{\beta}$ *under* (1.2) *can uniquely be written as*

$$\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = (\mathbf{K}\mathbf{A}^+ - \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+)\mathbf{b} + \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{y}, \tag{4.28}$$

$$\mathrm{E[\,OLSE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{K}\boldsymbol{\beta}, \quad \mathrm{D[\,OLSE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\,] = \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+)'. \tag{4.29}$$

(iii) *The OLSE of* $\mathbf{X}\boldsymbol{\beta}$ *under* (1.2) *is*

$$\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}) = (\mathbf{X}\mathbf{A}^+ - \widehat{\mathbf{X}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+)\mathbf{b} + \widehat{\mathbf{X}}\widehat{\mathbf{X}}^+\mathbf{y}, \tag{4.30}$$

$$\mathrm{E[\,OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\,] = \mathbf{X}\boldsymbol{\beta}, \quad \mathrm{D[\,OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\,] = \widehat{\mathbf{X}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}\widehat{\mathbf{X}}\widehat{\mathbf{X}}^+. \tag{4.31}$$

(iv) *The OLSE of* $\widetilde{\mathbf{X}}\boldsymbol{\beta}$ *under* (1.2) *is*

$$\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}) = (\widetilde{\mathbf{X}}\mathbf{A}^+ - \widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+)\mathbf{b} + \widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+\mathbf{y}, \tag{4.32}$$

$$\mathrm{E[\,OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})\,] = \widetilde{\mathbf{X}}\boldsymbol{\beta}, \quad \mathrm{D[\,OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})\,] = \widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}(\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+)'. \tag{4.33}$$

(v) *Under the conditions that* $\mathbf{X}_i\boldsymbol{\beta}_i$ *are estimable under* (1.2), *the OLSEs of* $\mathbf{X}_i\boldsymbol{\beta}_i$ *under* (1.2) *are*

$$\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{Y}_i\boldsymbol{\beta}) = (\mathbf{Y}_i\mathbf{A}^+ - \widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+)\mathbf{b} + \widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\mathbf{y}, \tag{4.34}$$

$$\mathrm{E[\,OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)\,] = \mathbf{X}_i\boldsymbol{\beta}_i, \quad \mathrm{D[\,OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)\,] = \widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}(\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+)' \tag{4.35}$$

    for $i = 1, 2, \ldots, k$.

(vi) *Under the conditions that* $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ *are estimable under* (1.2), *the OLSEs of* $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ *are*

$$\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i) = \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{W}_i\boldsymbol{\beta}) = (\mathbf{W}_i\mathbf{A}^+ - \widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+)\mathbf{b} + \widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\mathbf{y}, \tag{4.36}$$

$$\mathrm{E[\,OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)\,] = \widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i, \quad \mathrm{D[\,OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)\,] = \widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}(\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+)' \tag{4.37}$$

    for $i = 1, 2, \ldots, k$.

The following group of results on closed-form formulas of BLUEs and their properties are well known; see, e.g., [16, 42, 45].

LEMMA 4.5. *Let* $\mathbf{K} \in \mathbb{R}^{t \times p}$ *be given and assume that* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.1). *Then*

$$\mathrm{D}(\mathbf{L}\mathbf{y} - \mathbf{K}\boldsymbol{\beta}) = \min \quad s.t. \quad \mathrm{E}(\mathbf{L}\mathbf{y} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0} \Leftrightarrow \mathbf{L}[\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^\perp\,] = [\,\mathbf{K},\,\mathbf{0}\,]. \tag{4.38}$$

The matrix equation on the right-hand side is consistent, i.e.,

$$[\,\mathbf{K},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^\perp\,]^+[\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^\perp\,] = [\,\mathbf{K},\,\mathbf{0}\,]$$

holds under $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$, *while the general expression of* $\mathbf{L}$, *denoted by* $\mathbf{P}_{\mathbf{K};\mathbf{X};\boldsymbol{\Sigma}}$, *and the corresponding* $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ *can be written as*

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{K};\mathbf{X};\boldsymbol{\Sigma}}\mathbf{y} = \left([\,\mathbf{K},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^\perp\,]^+ + \mathbf{U}[\,\mathbf{X},\,\boldsymbol{\Sigma}\,]^\perp\right)\mathbf{y}, \tag{4.39}$$

$$\mathrm{E[BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})] = \mathbf{K}\boldsymbol{\beta}, \tag{4.40}$$

$$\mathrm{D[BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})] = [\,\mathbf{K},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^\perp\,]^+\boldsymbol{\Sigma}([\,\mathbf{K},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^\perp\,]^+)', \tag{4.41}$$

where $\mathbf{U} \in \mathbb{R}^{t \times n}$ *is arbitrary. In particular,*

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{X};\boldsymbol{\Sigma}}\mathbf{y} = \left([\,\mathbf{X},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^{\perp}\,]^{+} + \mathbf{U}[\,\mathbf{X},\,\boldsymbol{\Sigma}\,]^{\perp}\right)\mathbf{y}, \tag{4.42}$$

$$\mathrm{E}[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta}, \tag{4.43}$$

$$\mathrm{D}[\mathrm{BLUE}_{\mathscr{M}}(\mathbf{X}\boldsymbol{\beta})] = [\,\mathbf{X},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^{\perp}\,]^{+}\boldsymbol{\Sigma}([\,\mathbf{X},\,\mathbf{0}\,][\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^{\perp}\,]^{+})', \tag{4.44}$$

where $\mathbf{U} \in \mathbb{R}^{n \times n}$ *is arbitrary. Further, the following results hold.*

(i) $r[\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^{\perp}\,] = r[\,\mathbf{X},\,\boldsymbol{\Sigma}\,]$, $\mathscr{R}[\,\mathbf{X},\,\boldsymbol{\Sigma}\mathbf{X}^{\perp}\,] = \mathscr{R}[\,\mathbf{X},\,\boldsymbol{\Sigma}\,]$, *and* $\mathscr{R}(\mathbf{X}) \cap \mathscr{R}(\boldsymbol{\Sigma}\mathbf{X}^{\perp}) = \{\mathbf{0}\}$.

(ii) $\mathbf{P}_{\mathbf{K};\mathbf{X};\boldsymbol{\Sigma}}$ *is unique if and only if* $r[\,\mathbf{X},\,\boldsymbol{\Sigma}\,] = n$.

(iii) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ *is unique with probability 1 if and only if* $\mathscr{M}$ *is consistent.*

The BLUE of $\widehat{\mathbf{K}}\boldsymbol{\gamma}$ under (4.25) can be derived from Lemma 4.5. Substituting it into (4.27) yields the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under (1.2) as follows.

LEMMA 4.6. *Let* $\widetilde{\mathscr{M}}$ *be as given in* (1.2) *and assume that* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.2). *Then, the following results hold.*

(i) *The BLUE of* $\mathbf{K}\boldsymbol{\beta}$ *is*

$$\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = (\,\mathbf{K} - \mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{X})\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{y}, \tag{4.45}$$

$$\mathrm{E}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})] = \mathbf{K}\boldsymbol{\beta}, \tag{4.46}$$

$$\mathrm{D}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})] = \left([\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\right)\boldsymbol{\Sigma}\left([\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\right)', \tag{4.47}$$

where $\mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}} = [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+} + \mathbf{U}[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{\perp}$, *in which* $\mathbf{U} \in \mathbb{R}^{t \times n}$ *is arbitrary.*

(ii) *The BLUE of* $\mathbf{X}\boldsymbol{\beta}$ *is*

$$\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}) = (\,\mathbf{I}_{n} - \mathbf{P}_{\widehat{\mathbf{X}};\boldsymbol{\Sigma}})\mathbf{X}\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{y}, \tag{4.48}$$

$$\mathrm{E}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{X}\boldsymbol{\beta}, \tag{4.49}$$

$$\mathrm{D}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}([\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+})', \tag{4.50}$$

where $\mathbf{P}_{\widehat{\mathbf{X}};\boldsymbol{\Sigma}} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+} + \mathbf{U}[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{\perp}$, *in which* $\mathbf{U} \in \mathbb{R}^{n \times n}$ *is arbitrary.*

(iii) *The BLUE of* $\widetilde{\mathbf{X}}\boldsymbol{\beta}$ *is*

$$\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}) = (\,\widetilde{\mathbf{X}} - \mathbf{P}_{\widehat{\mathbf{Z}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{X})\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\widehat{\mathbf{Z}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{y}, \tag{4.51}$$

$$\mathrm{E}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})] = \widetilde{\mathbf{X}}\boldsymbol{\beta}, \tag{4.52}$$

$$\mathrm{D}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})] = [\,\widehat{\mathbf{Z}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}([\,\widehat{\mathbf{Z}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+})', \tag{4.53}$$

where $\mathbf{P}_{\widehat{\mathbf{Z}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}} = [\,\widehat{\mathbf{Z}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+} + \mathbf{V}[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{\perp}$, *and* $\mathbf{V} \in \mathbb{R}^{(m+n) \times n}$ *is arbitrary.*

(iv) *Under the conditions that* $\mathbf{X}_{i}\boldsymbol{\beta}_{i}$ *are estimable under* (1.2), *the BLUEs of* $\mathbf{X}_{i}\boldsymbol{\beta}_{i}$ *are*

$$\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_{i}\boldsymbol{\beta}_{i}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{Y}_{i}\boldsymbol{\beta}) = (\,\mathbf{Y}_{i} - \mathbf{P}_{\widehat{\mathbf{Y}}_{i};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{X})\mathbf{A}^{+}\mathbf{b} + \mathbf{P}_{\widehat{\mathbf{Y}}_{i};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{y},$$

$$\mathrm{E}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_{i}\boldsymbol{\beta}_{i})] = \mathbf{X}_{i}\boldsymbol{\beta}_{i},$$

$$\mathrm{D}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}_{i}\boldsymbol{\beta}_{i})] = [\,\widehat{\mathbf{Y}}_{i},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}([\,\widehat{\mathbf{Y}}_{i},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+})',$$

where $\mathbf{P}_{\widehat{\mathbf{Y}}_{i};\widehat{\mathbf{X}};\boldsymbol{\Sigma}} = [\,\widehat{\mathbf{Y}}_{i},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+} + \mathbf{U}_{i}[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{\perp}$ *and* $\mathbf{U}_{i} \in \mathbb{R}^{n \times n}$ *is arbitrary,* $i = 1, \ldots, k$.

(v) *Under the conditions that* $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ *are estimable under* (1.2)*, the BLUEs of* $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ *are*

$$\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{W}_i\boldsymbol{\beta}) = (\,\mathbf{W}_i - \mathbf{P}_{\widehat{\mathbf{W}}_i;\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{X})\mathbf{A}^+\mathbf{b} + \mathbf{P}_{\widehat{\mathbf{W}}_i;\widehat{\mathbf{X}};\boldsymbol{\Sigma}}\mathbf{y},$$
$$\mathrm{E}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)] = \widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i,$$
$$\mathrm{D}[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)] = [\,\widehat{\mathbf{W}}_i,\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\boldsymbol{\Sigma}([\,\widehat{\mathbf{W}}_i,\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+)',$$

where $\mathbf{P}_{\widehat{\mathbf{W}}_i;\widehat{\mathbf{X}};\boldsymbol{\Sigma}} = [\,\widehat{\mathbf{W}}_i,\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+ + \mathbf{U}_i[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^\perp,$ *and* $\mathbf{U}_i \in \mathbb{R}^{(m+n)\times n}$ *is arbitrary,* $i = 1,\dots,k$.

(vi) $r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,] = r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\,],\ \mathscr{R}[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,] = \mathscr{R}[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\,],$ *and* $\mathscr{R}(\widehat{\mathbf{X}}) \cap \mathscr{R}(\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp) = \{\mathbf{0}\}$.

(vii) $\mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\boldsymbol{\Sigma}}$ *is unique if and only if* $r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\,] = n$.

(viii) $\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ *is unique with probability* 1 *if and only if* $\widetilde{\mathscr{M}}$ *is consistent.*

5. Algebraic and statistical characterizations of equalities of OLSEs and BLUEs under CGLMs. Because OLSEs and BLUEs under CGLMs are defined from two different optimality criteria, they have different expressions and performances in the statistical inference of CGLMs. In fact, the lemmas in the previous section show that OLSEs and BLUEs of parameter spaces in CGLMs can be written as certain algebraic expressions composed by the observed response vectors, the given model matrices, and the covariance matrices of the error terms in CGLMs. It is advantageous to obtain these exact algebraic formulas of OLSEs and BLUEs of parameter spaces in CGLMs in the last section. From these exact algebraic expressions of OLSEs and BLUEs, as well as various algebraic tools in matrix theory, people have derived many valuable properties and features of OLSEs and BLUEs in the statistical literature, and have established a systematic theory on OLSEs and BLUEs and their applications. Even so, there are many new problems on OLSEs and BLUEs that can be proposed and approached from theoretical and applied points of view. Especially, we are more interested in describing algebraic and statistical properties and features of the estimators under various assumptions, and establishing relations between the OLSEs and BLUEs. Note from Lemmas 4.4 and 4.6 that OLSEs and BLUEs under CGLMs all can be represented by the given vectors and matrices in the models and their generalized inverses. Thus, to characterize the connections between OLSEs and BLUEs is in fact to characterize matrix equalities associated with OLSEs and BLUEs. In this case, we can use the formulas in Section 2 to establish and simplify various matrix expressions and matrix equalities related to OLSEs and BLUEs under CGLMs.

Classic results on the equalities of OLSEs and BLUEs under general linear statistical models were widely spread in the statistical literature. A group of latest results on the equality of the OLSE and BLUE of an estimable parametric function $\mathbf{K}\boldsymbol{\beta}$ under (1.1) were collected or proved in [65]. As an extension, we derive some necessary and sufficient conditions for the OLSE of $\mathbf{K}\boldsymbol{\beta}$ to be the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under (1.2).

THEOREM 5.1. *Let* $\mathbf{K} \in \mathbb{R}^{t\times p}$ *be given and assume that* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (1.2)*. Then, the following statistical and algebraic statements are equivalent:*

(i) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ *holds definitely.*

(ii) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ *holds with probability* 1.

(iii) $\mathrm{D}\left[\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\right] = \mathrm{D}\left[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\right].$

(iv) $\mathrm{Cov}\left\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}),\mathbf{y}\right\} = \mathrm{Cov}\left\{\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}),\mathbf{y}\right\}.$

(v) $\mathrm{Cov}\left\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}),\mathbf{y}\right\} = \mathrm{Cov}\left\{\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}),\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\right\}.$

(vi) $\mathrm{Cov}\left\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}),\mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\right\} = \mathbf{0}.$

(vii) $\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+ = [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+.$

(viii) $\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma} = [\,\widehat{\mathbf{K}},\, \mathbf{0}\,][\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\mathbf{\Sigma}$.

(ix) $\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma} = [\,\widehat{\mathbf{K}},\, \mathbf{0}\,][\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\mathbf{\Sigma}\widehat{\mathbf{X}}\widehat{\mathbf{X}}^+$.

(x) $\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+)' = [\,\widehat{\mathbf{K}},\, \mathbf{0}\,][\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\mathbf{\Sigma}([\,\widehat{\mathbf{K}},\, \mathbf{0}\,][\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+)'$.

(xi) $\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp = \mathbf{0}$.

(xii) $\widehat{\mathbf{X}}^\perp[\mathbf{\Sigma}\widehat{\mathbf{X}}, \mathbf{0}]\mathbf{F}_{[\widehat{\mathbf{X}}'\widehat{\mathbf{X}}, \widehat{\mathbf{K}}']} = \mathbf{0}$.

(xiii) $r\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{K}}' \\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \end{bmatrix} = 2r(\widehat{\mathbf{X}})$.

(xiv) $\mathscr{R}\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} \end{bmatrix} \supseteq \mathscr{R}\begin{bmatrix} \widehat{\mathbf{K}}' \\ \mathbf{0} \end{bmatrix}$.

(xv) $r\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{A}' & \mathbf{K}' \\ \mathbf{\Sigma}\mathbf{X} & \mathbf{X} & \mathbf{0} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} & \mathbf{0} \end{bmatrix} = 2r\begin{bmatrix} \mathbf{X} \\ \mathbf{A} \end{bmatrix} + r(\mathbf{A})$.

(xvi) $\mathscr{R}[(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma})'] \subseteq \mathscr{R}(\widehat{\mathbf{X}})$.

(xvii) $\mathscr{R}\begin{bmatrix} \widehat{\mathbf{X}}'\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp \\ \mathbf{0} \end{bmatrix} \subseteq \mathscr{R}\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} \\ \widehat{\mathbf{K}} \end{bmatrix}$.

(xviii) $\mathscr{R}\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{A}' \\ \mathbf{\Sigma}\mathbf{X} & \mathbf{X} & \mathbf{0} \\ \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \end{bmatrix} \supseteq \mathscr{R}\begin{bmatrix} \mathbf{K}' \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$.

Proof. It is easy to see from (4.28), (4.45), and Definition 1.1(i) that the equality $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ holds definitely if and only if the constant vectors and the coefficient matrices in (4.28) and (4.45) are equal, respectively, i.e.,

$$(\mathbf{K}\mathbf{A}^+ - \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+)\mathbf{b} = (\mathbf{K} - \mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\mathbf{\Sigma}}\mathbf{X})\mathbf{A}^+\mathbf{b} \quad \text{and} \quad \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+ = \mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\mathbf{\Sigma}}. \tag{5.54}$$

Combining the two equalities in (5.54) yields

$$\mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\mathbf{\Sigma}}[\mathbf{I}_n,\, \mathbf{X}\mathbf{A}^+\mathbf{b}] = [\,\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+,\, \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+\mathbf{b}\,]. \tag{5.55}$$

Substituting $\mathbf{P}_{\widehat{\mathbf{K}};\widehat{\mathbf{X}};\mathbf{\Sigma}}$ into (5.55) yields a matrix equation

$$\begin{aligned}
&\mathbf{U}[\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^\perp[\mathbf{I}_n,\, \mathbf{X}\mathbf{A}^+\mathbf{b}] \\
&= [\,\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+,\, \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+\mathbf{b}\,] - [\,\widehat{\mathbf{K}},\, \mathbf{0}\,][\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+[\mathbf{I}_n,\, \mathbf{X}\mathbf{A}^+\mathbf{b}].
\end{aligned} \tag{5.56}$$

From Lemma 2.3, the matrix equation in (5.56) is solvable for $\mathbf{U}$ if and only if

$$\begin{aligned}
&r\begin{bmatrix} [\,\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+,\, \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+\mathbf{b}\,] - [\,\widehat{\mathbf{K}},\, \mathbf{0}\,][\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+[\mathbf{I}_n,\, \mathbf{X}\mathbf{A}^+\mathbf{b}] \\ [\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^\perp[\mathbf{I}_n,\, \mathbf{X}\mathbf{A}^+\mathbf{b}] \end{bmatrix} \\
&= r\Big([\,\widehat{\mathbf{X}},\, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\,]^\perp[\mathbf{I}_n,\, \mathbf{X}\mathbf{A}^+\mathbf{b}]\Big),
\end{aligned} \tag{5.57}$$

where by Lemma 2.1, (2.6), (2.7), (2.9), and elementary block matrix operations, the both sides of (5.57)

reduce to

$$r\begin{bmatrix}[\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+,\ \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+\mathbf{b}]-[\widehat{\mathbf{K}},\ \mathbf{0}][\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+[\mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}]\\ [\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^\perp[\mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}]\end{bmatrix}$$

$$=r\begin{bmatrix}[\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+,\ \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+\mathbf{b}]-[\widehat{\mathbf{K}},\ \mathbf{0}][\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+[\mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}] & \mathbf{0}\\ [\mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}] & [\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]\end{bmatrix}$$

$$-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]$$

$$=r\begin{bmatrix}[\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+,\ \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{X}\mathbf{A}^+\mathbf{b}] & [\widehat{\mathbf{K}},\ \mathbf{0}]\\ [\mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}] & [\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]\end{bmatrix}-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]$$

$$=r\begin{bmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix}-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]$$

$$=n+r(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp)-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]$$

$$=r\begin{bmatrix}\widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \widehat{\mathbf{X}}'\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp\\ \widehat{\mathbf{K}} & \mathbf{0}\end{bmatrix}-r(\widehat{\mathbf{X}})+n-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]\ \ \text{(by (2.9))}$$

$$=r\begin{bmatrix}\widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{K}}'\\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0}\end{bmatrix}-2r(\widehat{\mathbf{X}})+n-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]\ \ \text{(by (2.7))}$$

$$=r\begin{bmatrix}\mathbf{F}_{\mathbf{A}}\mathbf{X}'\mathbf{X}\mathbf{F}_{\mathbf{A}} & \mathbf{0} & \mathbf{F}_{\mathbf{A}}\mathbf{K}'\\ \mathbf{\Sigma}\mathbf{X}\mathbf{F}_{\mathbf{A}} & \mathbf{X}\mathbf{F}_{\mathbf{A}} & \mathbf{0}\end{bmatrix}-2r(\mathbf{X}\mathbf{F}_{\mathbf{A}})+n-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]$$

$$=r\begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{A}' & \mathbf{K}'\\ \mathbf{\Sigma}\mathbf{X} & \mathbf{X} & \mathbf{0} & \mathbf{0}\\ \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{A} & \mathbf{0} & \mathbf{0}\end{bmatrix}-2r\begin{bmatrix}\mathbf{X}\\ \mathbf{A}\end{bmatrix}-r(\mathbf{A})+n-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}]\ \ \text{(by (2.6) and (2.7)),}\qquad(5.58)$$

and

$$r\left([\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^\perp[\mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}]\right)$$
$$=r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp,\ \mathbf{I}_n,\ \mathbf{X}\mathbf{A}^+\mathbf{b}]-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]=n-r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}].\qquad(5.59)$$

Substituting (5.58) and (5.59) into (5.57), we obtain the following equalities

$$r(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp)=\mathbf{0},\quad r\begin{bmatrix}\widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{K}}'\\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0}\end{bmatrix}=2r(\widehat{\mathbf{X}}),$$

$$r\begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{0} & \mathbf{A}' & \mathbf{K}'\\ \mathbf{\Sigma}\mathbf{X} & \mathbf{X} & \mathbf{0} & \mathbf{0}\\ \mathbf{A} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{A} & \mathbf{0} & \mathbf{0}\end{bmatrix}=2r\begin{bmatrix}\mathbf{X}\\ \mathbf{A}\end{bmatrix}+r(\mathbf{A}),$$

so that (i) and (xi)–(xviii) are equivalent.

Since $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ and $\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ are both unbiased for $\mathbf{K}\boldsymbol{\beta}$, then we obtain from Definition 1.1(ii) that $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})$ holds with probability 1 if and only if $\mathrm{D}[\,\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})-\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\,]=$

$\mathbf{0}$. From (4.28) and (4.45),

$$
\begin{aligned}
&\mathrm{D}[\,\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) - \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\,] \\
&= \left(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+ - [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\right)\boldsymbol{\Sigma}\left(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+ - [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\right)'.
\end{aligned}
$$

Hence, $\mathrm{D}[\,\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta}) - \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})\,] = \mathbf{0}$ is equivalent to the matrix equality

$$
\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma} = [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\boldsymbol{\Sigma},
$$

where by (2.9) and elementary block matrix operations,

$$
\begin{aligned}
r\left(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma} - [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\boldsymbol{\Sigma}\right) &= r\begin{bmatrix} [\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,] & \boldsymbol{\Sigma} \\ [\,\widehat{\mathbf{K}},\,\mathbf{0}\,] & \widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma} \end{bmatrix} - r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,] \\
&= r\begin{bmatrix} [\,\widehat{\mathbf{X}},\,\mathbf{0}\,] & \boldsymbol{\Sigma} \\ [\,\mathbf{0},\,-\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,] & \mathbf{0} \end{bmatrix} - r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\,] \\
&= r(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^+\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp).
\end{aligned}
\tag{5.60}
$$

Setting the right-hand side of (5.60) equal to zero leads to the equivalences of (ii), (viii), and (xi).

Note that

$$
\mathscr{R}(\widehat{\mathbf{K}}') \subseteq \mathscr{R}(\widehat{\mathbf{X}}'), \quad \mathscr{R}([\,\widehat{\mathbf{K}},\,\mathbf{0}\,]') \subseteq \mathscr{R}([\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]'), \quad \mathscr{R}(\boldsymbol{\Sigma}) \subseteq \mathscr{R}([\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]).
$$

Applying (2.10) to (4.29) and (4.47), and simplifying by (2.8) and elementary block matrix operations, we obtain

$$
\begin{aligned}
&r\big(\mathrm{D}[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})] - \mathrm{D}[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{K}\boldsymbol{\beta})]\big) \\
&= r\Big(\widehat{\mathbf{K}}(\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^+\widehat{\mathbf{X}}'\boldsymbol{\Sigma}\widehat{\mathbf{X}}(\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^+\widehat{\mathbf{K}}' - [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+\boldsymbol{\Sigma}([\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]^+)'[\,\widehat{\mathbf{K}},\,\mathbf{0}\,]'\Big) \\
&= r\left([\,\widehat{\mathbf{K}},\,\widehat{\mathbf{K}},\,\mathbf{0}\,]\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & [\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,] \end{bmatrix}^+\begin{bmatrix} \widehat{\mathbf{X}}'\boldsymbol{\Sigma}\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{\Sigma} \end{bmatrix}\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & [\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^\perp\,]' \end{bmatrix}^+\begin{bmatrix} \widehat{\mathbf{K}}' \\ \widehat{\mathbf{K}}' \\ \mathbf{0} \end{bmatrix}\right)
\end{aligned}
$$

$$= r\begin{bmatrix} \mathbf{0} & \begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & [\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]' \end{bmatrix} & \begin{bmatrix} \widehat{\mathbf{K}}' \\ \widehat{\mathbf{K}}' \\ \mathbf{0} \end{bmatrix} \\ \begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & [\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}] \end{bmatrix} & \begin{bmatrix} \widehat{\mathbf{X}}'\mathbf{\Sigma}\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{\Sigma} \end{bmatrix} & \mathbf{0} \\ [\widehat{\mathbf{K}}, [\widehat{\mathbf{K}}, \mathbf{0}]] & \mathbf{0} & \mathbf{0} \end{bmatrix} - 2r(\widehat{\mathbf{X}}) - 2r[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]$$

$$= r\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{K}}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}' & \widehat{\mathbf{K}}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma} & \mathbf{0} \\ \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}'\mathbf{\Sigma}\widehat{\mathbf{X}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{X}} & \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp} & \mathbf{0} & -\mathbf{\Sigma} & \mathbf{0} \\ \widehat{\mathbf{K}} & \widehat{\mathbf{K}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} - 2r(\widehat{\mathbf{X}}) - 2r[\widehat{\mathbf{X}}, \mathbf{\Sigma}]$$

$$= r\begin{bmatrix} \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & -\widehat{\mathbf{X}}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}' & \widehat{\mathbf{K}}' \\ \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{X}}'\mathbf{\Sigma}\widehat{\mathbf{X}} & \mathbf{0} & \mathbf{0} \\ -\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} & -\mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{K}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} + r(\widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}) - 2r(\widehat{\mathbf{X}}) - 2r[\widehat{\mathbf{X}}, \mathbf{\Sigma}]$$

$$= r\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} & \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{K}}' \\ \widehat{\mathbf{X}}'\mathbf{\Sigma} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \widehat{\mathbf{X}}' & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{K}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} - 3r(\widehat{\mathbf{X}}) - r[\widehat{\mathbf{X}}, \mathbf{\Sigma}]$$

$$= r\begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} & \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{K}}' \end{bmatrix} + r\begin{bmatrix} \widehat{\mathbf{X}}'\mathbf{\Sigma} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} \\ \widehat{\mathbf{X}}' & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{K}} \end{bmatrix} - 3r(\widehat{\mathbf{X}}) - r[\widehat{\mathbf{X}}, \mathbf{\Sigma}] \text{ (by (2.8))}$$

$$= r\begin{bmatrix} \widehat{\mathbf{X}}'\mathbf{\Sigma} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} \\ \widehat{\mathbf{X}}' & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{K}} \end{bmatrix} - 2r(\widehat{\mathbf{X}}). \tag{5.61}$$

Setting the right-hand side of (5.61) equal to zero leads to the equivalences of (iii), (x), and (xiii).

$$r\left(\widehat{\mathbf{K}}\widehat{\mathbf{X}}^{+} - [\widehat{\mathbf{K}}, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]^{+}\right)$$

$$= r\left(\widehat{\mathbf{K}}(\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^{+}\widehat{\mathbf{X}}' - [\widehat{\mathbf{K}}, \mathbf{0}]([\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]'[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}])^{+}[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]'\right)$$

$$= r\left[(\widehat{\mathbf{K}}, [\widehat{\mathbf{K}}, \mathbf{0}])\begin{pmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} \\ \mathbf{0} & -[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]'[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}] \end{pmatrix}^{+}\begin{bmatrix} \widehat{\mathbf{X}}' \\ [\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]' \end{bmatrix}\right]$$

$$= r\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{X}}' \\ \mathbf{0} & -[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]'[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}] & [\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]' \\ \widehat{\mathbf{K}} & [\widehat{\mathbf{K}}, \mathbf{0}] & \mathbf{0} \end{bmatrix} - r[\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}] - r(\widehat{\mathbf{X}})$$

$$= r \begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \widehat{\mathbf{X}}'\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp} & \widehat{\mathbf{X}}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}^{\perp}\boldsymbol{\Sigma} \\ \widehat{\mathbf{K}} & \widehat{\mathbf{K}} & \mathbf{0} & \mathbf{0} \end{bmatrix} - r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\,] - r(\widehat{\mathbf{X}})$$

$$= r \begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \widehat{\mathbf{X}}'\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}' \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{X}}^{\perp}\boldsymbol{\Sigma} \\ \widehat{\mathbf{K}} & \mathbf{0} & \mathbf{0} \end{bmatrix} - r[\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\,] - r(\widehat{\mathbf{X}})$$

$$= r \begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \widehat{\mathbf{X}}'\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp} \\ \widehat{\mathbf{K}} & \mathbf{0} \end{bmatrix} - r(\widehat{\mathbf{X}}) = r \begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \widehat{\mathbf{X}}'\boldsymbol{\Sigma} \\ \widehat{\mathbf{K}} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{X}}' \end{bmatrix} - 2r(\widehat{\mathbf{X}}). \tag{5.62}$$

Setting the right-hand side of (5.62) equal to zero leads to the equivalence of (vii) and (xiii).

Note that

$$[\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp} = \mathbf{0} \Leftrightarrow [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma} = [\,\widehat{\mathbf{K}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}\widehat{\mathbf{X}}\widehat{\mathbf{X}}^{+}.$$

Hence, (viii) and (ix) are equivalent. The equivalences of (iv) and (viii), (v) and (ix), (vi) and (xi) follow from the corresponding covariance matrix calculations.   □

Setting $\mathbf{K} = \mathbf{X}$ and $\mathbf{K} = \widetilde{\mathbf{X}}$ in Theorem 5.1, respectively, we obtain the following consequences.

COROLLARY 5.2. *Let* $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})$ *and* $\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})$ *be as given in* (4.30) *and* (4.48), *respectively. Then, the following statements are equivalent*:

(i) $\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})$ *holds definitely* (*with probability* 1).

(ii) $\mathrm{D}\big[\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big] = \mathrm{D}\big[\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big]$.

(iii) $\mathrm{D}[\mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] = \mathrm{D}[\mathbf{y} - \mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})]$.

(iv) $\mathrm{D}(\mathbf{y}) = \mathrm{D}[\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] + \mathrm{D}[\mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})]$.

(v) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\}$.

(vi) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.

(vii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\}$.

(viii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.

(ix) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathbf{y}, \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.

(x) $\mathrm{Cov}\{\mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathbf{y}, \mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.

(xi) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$
     $= \mathrm{Cov}\{\mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.

(xii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} + \mathrm{Cov}\{\mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathbf{0}$.

(xiii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widetilde{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathbf{0}$.

(xiv) $\mathbf{P}_{\widehat{\mathbf{X}}} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}$.

(xv) $\mathbf{P}_{\widehat{\mathbf{X}}}\boldsymbol{\Sigma} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}$.

(xvi) $\mathbf{P}_{\widehat{\mathbf{X}}}\boldsymbol{\Sigma} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$.

(xvii) $\mathbf{P}_{\widehat{\mathbf{X}}}\boldsymbol{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}([\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+})'$.

(xviii) $\mathbf{P}_{\widehat{\mathbf{X}}}\boldsymbol{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}$.

(xix) $\mathbf{P}_{\widehat{\mathbf{X}}}\boldsymbol{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}} = [\,\widehat{\mathbf{X}},\,\mathbf{0}\,][\,\widehat{\mathbf{X}},\,\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp}\,]^{+}\boldsymbol{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$.

(xx) $\mathbf{P}_{\widehat{\mathbf{X}}}\boldsymbol{\Sigma}\widehat{\mathbf{X}}^{\perp} = \mathbf{0}$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 32, pp. 232-253, July 2017.

Bo Jiang and Yongge Tian   248

(xxi) $\widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}} = \mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}$.

(xxii) $\widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}} + \mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp} = \mathbf{0}$.

(xxiii) $\mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma} = \mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$.

(xxiv) $\widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma} = \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}$.

(xxv) $r[\mathbf{\Sigma}\widehat{\mathbf{X}},\ \widehat{\mathbf{X}}] = r(\widehat{\mathbf{X}})$.

(xxvi) $r[\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp},\ \widehat{\mathbf{X}}^{\perp}] = r(\widehat{\mathbf{X}}^{\perp})$.

(xxvii) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}) \subseteq \mathscr{R}(\widehat{\mathbf{X}})$.

(xxviii) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}) \subseteq \mathscr{R}(\widehat{\mathbf{X}}^{\perp})$.

(xxix) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}) = \mathscr{R}(\mathbf{\Sigma}) \cap \mathscr{R}(\widehat{\mathbf{X}})$.

(xxx) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}) = \mathscr{R}(\mathbf{\Sigma}) \cap \mathscr{R}(\widehat{\mathbf{X}}^{\perp})$.

Proof. Results (i), (ii), (v), (vi), (xiii)–(xvii), (xx), (xxv) follow from Theorem 5.1. The proofs of the equivalences of (iii), (vii), (viii), (ix), (xviii), (xix), (xxiii)–(xxv) are similar to those presented in [65].

From (1.1) and (4.30),

$$\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}),\ \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\} + \mathrm{Cov}\big\{\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}),\ \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\}$$
$$= \mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp} + \widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}, \tag{5.63}$$

and

$$\mathrm{D}(\mathbf{y}) - \mathrm{D}[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] - \mathrm{D}[\mathbf{Y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] = \mathbf{\Sigma} - \mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}} - \widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}$$
$$= \mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp} + \widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}, \tag{5.64}$$

where by (2.6) and Lemma 2.1(v),

$$r\big(\mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp} + \widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}\big) = r\big(\mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}\big) + r\big(\widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}\big) = 2r\big(\widehat{\mathbf{X}}^{\perp}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}\big)$$
$$= 2r[\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}] - 2r(\widehat{\mathbf{X}}). \tag{5.65}$$

Setting both sides of (5.65) equal to zero and combining it with (5.63) and (5.64) yield the equivalences of (iv), (xii), (xxi), (xxii), (xxv), and (xxvii).

The equivalences of (xxiii)–(xxx) on matrix equalities and range equalities are well known; see [41]. The equivalence of (x) and (xxiv), and the equivalence of (xi) and (xxi) follow from the formulas for the corresponding covariance matrix operations. □

COROLLARY 5.3. *Let* $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})$ *and* $\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})$ *be as given in* (4.32) *and* (4.51), *respectively. Then, the following statements are equivalent*:

(i) $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})$ *holds definitely* (*with probability 1*).

(ii) $\mathrm{D}\big[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})\big] = \mathrm{D}\big[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})\big]$.

(iii) $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}),\mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}),\mathbf{y}\big\}$.

(iv) $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}),\mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}),\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\}$.

(v) $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}),\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\} = \mathbf{0}$.

(vi) $\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^{+} = [\widehat{\mathbf{Z}},\ \mathbf{0}][\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]^{+}$.

(vii) $\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^{+}\mathbf{\Sigma} = [\widehat{\mathbf{Z}},\ \mathbf{0}][\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]^{+}\mathbf{\Sigma}$.

(viii) $\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^{+}\mathbf{\Sigma} = [\widehat{\mathbf{Z}},\ \mathbf{0}][\widehat{\mathbf{X}},\ \mathbf{\Sigma}\widehat{\mathbf{X}}^{\perp}]^{+}\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$.

(ix) $\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}(\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+)' = [\widehat{\mathbf{Z}}, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}([\widehat{\mathbf{Z}}, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+)'$.

(x) $\widehat{\mathbf{Z}}\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp = \mathbf{0}$.

(xi) $\mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma} = \mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$.

(xii) $\widehat{\mathbf{X}}^\perp\mathbf{\Sigma} = \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp$.

(xiii) $r[\mathbf{\Sigma}\widehat{\mathbf{X}}, \widehat{\mathbf{X}}] = r(\widehat{\mathbf{X}})$.

(xiv) $r[\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp, \widehat{\mathbf{X}}^\perp] = r(\widehat{\mathbf{X}}^\perp)$.

(xv) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}) \subseteq \mathscr{R}(\widehat{\mathbf{X}})$.

(xvi) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp) \subseteq \mathscr{R}(\widehat{\mathbf{X}}^\perp)$.

(xvii) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}) = \mathscr{R}(\mathbf{\Sigma}) \cap \mathscr{R}(\widehat{\mathbf{X}})$.

(xviii) $\mathscr{R}(\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp) = \mathscr{R}(\mathbf{\Sigma}) \cap \mathscr{R}(\widehat{\mathbf{X}}^\perp)$.

Setting $\mathbf{K} = \mathbf{Y}_i$ and $\mathbf{K} = \mathbf{W}_i$, respectively, we obtain from Theorem 5.1 the following consequences.

COROLLARY 5.4. *Assume that all $\mathbf{X}_i\boldsymbol{\beta}_i$ and $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ are estimable under* (1.2). *Then, the following results hold.*

(a) *The following statements are equivalent*:

(i) $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)$ *holds definitely (with probability 1)*, $i = 1, \ldots, k$.

(ii) $\mathrm{D}[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)] = \mathrm{D}[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)]$, $i = 1, \ldots, k$.

(iii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y}\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y}\}$, $i = 1, \ldots, k$.

(iv) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y}\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$, $i = 1, \ldots, k$.

(v) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathbf{0}$, $i = 1, \ldots, k$.

(vi) $\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+ = [\widehat{\mathbf{Y}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+$, $i = 1, \ldots, k$.

(vii) $\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma} = [\widehat{\mathbf{Y}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}$, $i = 1, \ldots, k$.

(viii) $\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma} = [\widehat{\mathbf{Y}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$, $i = 1, \ldots, k$.

(ix) $\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma}(\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+)' = [\widehat{\mathbf{Y}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}([\widehat{\mathbf{Y}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+)'$, $i = 1, 2, \ldots, k$.

(x) $\widehat{\mathbf{Y}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp = \mathbf{0}$, $i = 1, \ldots, k$.

(xi) $\widehat{\mathbf{X}}^\perp[\mathbf{\Sigma}\widehat{\mathbf{X}}, \mathbf{0}]\mathbf{F}_{[\widehat{\mathbf{X}}'\widehat{\mathbf{X}}, \widehat{\mathbf{Y}}'_i]} = \mathbf{0}$, $i = 1, \ldots, k$.

(xii) $r\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{Y}}'_i \\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \widehat{\mathbf{Z}}'_i\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{Y}}'_i \\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \end{bmatrix} = 2r(\widehat{\mathbf{X}})$, $i = 1, \ldots, k$.

(b) *The following statements are equivalent*:

(i) $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)$ *holds definitely (with probability 1)*, $i = 1, \ldots, k$.

(ii) $\mathrm{D}\left[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)\right] = \mathrm{D}\left[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)\right]$, $i = 1, \ldots, k$.

(iii) $\mathrm{Cov}\left\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y}\right\} = \mathrm{Cov}\left\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y}\right\}$, $i = 1, \ldots, k$.

(iv) $\mathrm{Cov}\left\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y}\right\} = \mathrm{Cov}\left\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\right\}$, $i = 1, \ldots, k$.

(v) $\mathrm{Cov}\left\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\right\} = \mathbf{0}$, $i = 1, \ldots, k$.

(vi) $\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+ = [\widehat{\mathbf{W}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+$, $i = 1, \ldots, k$.

(vii) $\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma} = [\widehat{\mathbf{W}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}$, $i = 1, \ldots, k$.

(viii) $\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma} = [\widehat{\mathbf{W}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}\mathbf{P}_{\widehat{\mathbf{X}}}$, $i = 1, \ldots, k$.

(ix) $\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma}(\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+)' = [\widehat{\mathbf{W}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+\mathbf{\Sigma}([\widehat{\mathbf{W}}_i, \mathbf{0}][\widehat{\mathbf{X}}, \mathbf{\Sigma}\widehat{\mathbf{X}}^\perp]^+)'$, $i = 1, \ldots, k$.

(x) $\widehat{\mathbf{W}}_i\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp = \mathbf{0}$, $i = 1, \ldots, k$.

(xi) $r\begin{bmatrix} \widehat{\mathbf{X}}'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{W}}'_i \\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \end{bmatrix} = 2r(\widehat{\mathbf{X}})$, $i = 1, \ldots, k$.

Proof. The equivalences of (i)–(xii) in (a) and the equivalences of (i)–(xi) in (b) follow from Theorem 5.1. □

We conclude with a group of amusing and versatile results concerning the statistical interpretations of the equivalences of the OLSEs and BLUEs of the parameter spaces in (1.2).

THEOREM 5.5. *Let* $\widetilde{\mathscr{M}}$ *be as given in* (1.2), *and assume that all* $\mathbf{X}_i\boldsymbol{\beta}_i$ *and* $\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i$ *are estimable under* (1.2). *Then, the following statistical facts are equivalent:*

  (i) $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})$ *holds definitely (with probability 1).*
  (ii) $\mathrm{D}\big[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big] = \mathrm{D}\big[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big]$.
  (iii) $\mathrm{D}[\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] = \mathrm{D}[\mathbf{y} - \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})]$.
  (iv) $\mathrm{D}(\mathbf{y}) = \mathrm{D}[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})] + \mathrm{D}[\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})]$.
  (v) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\}$.
  (vi) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.
  (vii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\}$.
  (viii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathrm{Cov}\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.
  (ix) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathbf{y}, \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.
  (x) $\mathrm{Cov}\{\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y}\} = \mathrm{Cov}\{\mathbf{y}, \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.
  (xi) $\mathrm{Cov}\{\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\}$.
  (xii) $\mathrm{Cov}\{\mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} + \mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathbf{0}$.
  (xiii) $\mathrm{Cov}\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\} = \mathbf{0}$.
  (xiv) $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})$ *holds definitely (with probability 1).*
  (xv) $\mathrm{D}\big[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})\big] = \mathrm{D}\big[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta})\big]$.
  (xvi) $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}), \mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}), \mathbf{y}\big\}$.
  (xvii) $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}), \mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\}$.
  (xviii) $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}\boldsymbol{\beta}), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\} = \mathbf{0}$.
  (xix) *All* $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)$ *hold definitely (with probability 1),* $i = 1, \ldots, k$.
  (xx) *All* $\mathrm{D}\big[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)\big] = \mathrm{D}\big[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i)\big]$ *hold,* $i = 1, \ldots, k$.
  (xxi) *All* $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y}\big\}$ *hold,* $i = 1, \ldots, k$.
  (xxii) *All* $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\}$ *hold,* $i = 1, \ldots, k$.
  (xxiii) *All* $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}_i\boldsymbol{\beta}_i), \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\} = \mathbf{0}$ *hold,* $i = 1, \ldots, k$.
  (xxiv) *All* $\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)$ *hold definitely (with probability 1),* $i = 1, \ldots, k$.
  (xxv) *All* $\mathrm{D}\big[\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)\big] = \mathrm{D}\big[\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i)\big]$ *hold,* $i = 1, \ldots, k$.
  (xxvi) *All* $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y}\big\}$ *hold,* $i = 1, \ldots, k$.
  (xxvii) *All* $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathbf{y}\big\} = \mathrm{Cov}\big\{\mathrm{BLUE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\}$ *hold,* $i = 1, \ldots, k$.
  (xxviii) *All* $\mathrm{Cov}\big\{\mathrm{OLSE}_{\widehat{\mathscr{M}}}(\widetilde{\mathbf{X}}_i\boldsymbol{\beta}_i), \ \mathbf{y} - \mathrm{OLSE}_{\widehat{\mathscr{M}}}(\mathbf{X}\boldsymbol{\beta})\big\} = \mathbf{0}$ *hold,* $i = 1, \ldots, k$.

Proof. The equivalences of (i)–(xiii) and the equivalences of (xiv)–(xviii) follow from Corollaries 5.2 and 5.3.

If (i) holds, then we derive from Corollary 5.2(xxvii) that

$$r\begin{bmatrix} \widehat{\mathbf{Z}}_i'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{Y}}_i' \\ \mathbf{\Sigma}\widehat{\mathbf{X}} & \widehat{\mathbf{X}} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \widehat{\mathbf{Z}}_i'\widehat{\mathbf{X}} & \mathbf{0} & \widehat{\mathbf{Y}}_i' \\ \mathbf{0} & \widehat{\mathbf{X}} & \mathbf{0} \end{bmatrix} = r[\widehat{\mathbf{Z}}_i'\widehat{\mathbf{X}},\ \widehat{\mathbf{Y}}_i'] + r(\widehat{\mathbf{X}})$$

$$= r[\widehat{\mathbf{X}}'\widehat{\mathbf{X}},\ \widehat{\mathbf{Y}}_i'] + r(\widehat{\mathbf{X}})$$

$$= r(\widehat{\mathbf{X}}'\widehat{\mathbf{X}}) + r(\widehat{\mathbf{X}})$$

$$= 2r(\widehat{\mathbf{X}}),\quad i = 1, \ldots, k,$$

thus establishing the equality(xii) in Corollary 5.4(a). Hence, (i) implies (xix)–(xxiii). If (xix) holds, we obtain from $\widehat{\mathbf{X}} = \widehat{\mathbf{Y}}_1 + \cdots + \widehat{\mathbf{Y}}_k$ and (x) in Corollary 5.4(a)

$$\mathbf{P}_{\widehat{\mathbf{X}}}\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp = \widehat{\mathbf{Y}}_1\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp + \cdots + \widehat{\mathbf{Y}}_k\widehat{\mathbf{X}}^+\mathbf{\Sigma}\widehat{\mathbf{X}}^\perp = \mathbf{0}.$$

Thus, Corollary 5.2(xx) holds.

From Corollaries 5.2 and 5.3, the equivalence of (i) and (xiv) holds obviously. The proof of the equivalence of (xiv) and (xxiv) is similar to that of the equivalence of (i) and (xix). □

6. Summary comments. We have approached OLSEs and BLUEs of whole and partial mean parameter vectors in CGLMs, and have established a variety of algebraic and statistical interpretations for the OLSEs to be the BLUEs of the parameter spaces in (1.2) by using some classic and novel algebraic tools and techniques in matrix analysis. All the results obtained demonstrate essential links between OLSEs and BLUEs under various assumptions, which thus enable us to recognize and use these equivalent statements in many different situations, and can serve as general references in the statistical inference of CGLMs. This work also shows that while even for some classic inference problems on linear statistical models, we are still able to reveal a variety of novel and insightful conclusions by making use of some effective matrix analysis tools. Thus, important developments of statistics always need essential supports from matrix theory and linear algebra. Finally, we mention that various similar problems on decompositions of OLSEs and BLUEs under linear statistical models, such as, linear models with both fixed- and random-effects, multivariate general linear models, etc., while the matrix methodology used in the previous sections is also available for approaching these problems under these general situations.

REFERENCES

[1]   A. Abdulle and G. Wanner. 200 years of least squares method. *Elem. Math.*, 57:45–60, 2002.
[2]   I.S. Alalouf and G.P.H. Styan. Characterizations of estimability in the general linear model. *Ann. Statist.*, 7:194–200, 1979.
[3]   I.S. Alalouf and G.P.H. Styan. Characterizations of the conditions for the ordinary least squares estimator to be best linear unbiased. In: Y.P. Chaubey and T.D. Dwivedi (editors), *Topics in Applied Statistics*, Department of of Mathematics, Concordia University, Montréal, 331–344, 1984.
[4]   J.K. Baksalary. Criteria for the equality between ordinary least squares and best linear unbiased estimators under certain linear models. *Canad. J. Stat.*, 16:97–102, 1988.

[5]   J.K. Baksalary and R. Kala. An extension of a rank criterion for the least squares estimator to be the best linear unbiased estimator. *J. Statist. Plann. Inference*, 1:309–312, 1977.

[6]   J.K. Baksalary and R. Kala. Simple least squares estimation versus best linear unbiased prediction. *J. Statist. Plann. Inference*, 5:147–151, 1981.

[7]   J.K. Baksalary and R. Kala. On equalities between BLUEs, WLSEs, and SLSEs. *Canad. J. Stat.*, 11:119–123, 1983 (Correction: 12:240, 1984).

[8]   J.K. Baksalary and P.R. Pordzik. Inverse-partitioned-matrix method for the general Gauss–Markov model with linear restrictions. *J. Statist. Plann. Inference*, 23:133–143, 1989.

[9]   J.K. Baksalary and P.R. Pordzik. Implied linear restrictions in the general Gauss–Markov model. *J. Statist. Plann. Inference*, 30:237–248, 1992.

[10] J.K. Baksalary and S. Puntanen. Characterizations of the best linear unbiased estimator in the general Gauss–Markov model with the use of matrix partial orderings. *Linear Algebra Appl.*, 127:363–370, 1990.

[11] J.K. Baksalary, S. Puntanen, and G.P.H. Styan. On T.W. Anderson's contributions to solving the problem of when the ordinary least-squares estimator is best linear unbiased and to characterizing rank additivity of matrices. In: G.P.H. Styan (editor), *The Collected Papers of T.W. Anderson: 1943–1985*, Wiley, New York, Vol. 2, 1579–1591, 1990.

[12] J.K. Baksalary and A.C. van Eijnsbergen. A comparison of two criteria for ordinary-least-squares estimators to be best linear unbiased estimators. *Amer. Statistican*, 42:205–208, 1988.

[13] O.M. Baksalary and G. Trenkler. Between OLSE and BLUE. *Aust. N.Z.J. Stat.*, 53:289–303, 2011.

[14] O.M. Baksalary, G. Trenkler, and E. Liski. Let us do the twist again. *Stat. Papers*, 54:1109–1119, 2013.

[15] G. Beganu. On some alternative forms equivalent to Kruskal's condition for OLSE to be BLUE. *Rev. R. Acad. Cien. Serie A. Mat.*, 101:159–166, 2007.

[16] H. Drygas. *The Coordinate-free Approach to Gauss–Markov Estimation*. Springer, Heidelberg, 1970.

[17] R.W. Farebrother. Some early statistical contributions to the theory and practice of linear algebra. *Linear Algebra Appl.*, 237/238:205–224, 1996.

[18] F.A. Graybill. *An Introduction to Linear Statistical Models*, Vol. I. McGraw–Hill, New York, 1961.

[19] S.J. Haberman. How much do Gauss–Markov and least squares estimators differ? A coordinate free approach. *Ann. Statist.*, 3:982–990, 1975.

[20] S.J. Haslett, J. Isotalo, Y. Liu, and S. Puntanen. Equalities between OLSE, BLUE and BLUP in the linear model. *Stat. Papers*, 55:543–561, 2014.

[21] S.J. Haslett and S. Puntanen. A note on the equality of the BLUPs for new observations under two linear models. *Acta Comment. Univ. Tartu. Math.*, 14:27–33, 2010.

[22] S.J. Haslett and S. Puntanen. Equality of BLUEs or BLUPs under two linear models using stochastic restrictions. *Stat. Papers*, 51:465–475, 2010.

[23] S.J. Haslett and S. Puntanen. On the equality of the BLUPs under two linear mixed models. *Metrika*, 74:381–395, 2011.

[24] R.R. Hocking. *The Analysis of Linear Models*. Brooks/Cole, Monterey, California, 1985.

[25] J. Isotalo and S. Puntanen. A note on the equality of the OLSE and the BLUE of the parametric functions in the general Gauss–Markov model. *Stat. Papers*, 50:185–193, 2009.

[26] A. Jaeger and W. Krämer. A final twist on the equality of OLS and GLS. *Stat. Papers*, 39:321–324, 1998.

[27] B. Jiang and Y. Sun. On the equality of estimators under a general partitioned linear model with parameter restrictions. *Stat. Papers*, DOI:10.1007/s00362-016-0837-9, 2017.

[28] W. Krämer. A note on the equality of ordinary least squares and Gauss–Markov estimates in the general linear model. *Sankhyā Ser. A*, 42:130–131, 1980.

[29] W. Krämer, R. Bartels, and D.G. Fiebig. Another twist on the equality of OLS and GLS. *Stat. Papers*, 37:277–281, 1996.

[30] W. Kruskal. When are Gauss–Markov and least squares estimators identical? A coordinate-free approach. *Ann. Math. Statist.*, 39:70–75, 1968.

[31] Y. Liu. On equality of ordinary least squares estimator, best linear unbiased estimator and best linear unbiased predictor in the general linear model. *J. Statist. Plann. Inference*, 139:1522–1529, 2009.

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

[33] F.W. McElroy. A necessary and sufficient condition that ordinary least-squares estimators be best linear unbiased. *J. Amer. Stat. Assoc.*, 62:1302–1304, 1967.

[34] G.A. Milliken. New criteria for estimability for linear models. *Ann. Math. Statist.*, 42:1588–1594, 1971.

[35] G.A. Milliken and M. Albohali. On necessary and sufficient conditions for ordinary least squares estimators to be best linear unbiased estimators. *Amer. Statistican*, 38:298–299, 1984.

[36] S.K. Mitra and B.J. Moore. Gauss–Markov estimation with an incorrect dispersion matrix. *Sankhyā Ser. A*, 35:139–152, 1973.

[37] H. Neudecker. Best linear unbiased estimation of β subject to linear equality constraints in the general linear model. *Stat. Papers*, 20:176–182, 1979.

[38] U. Norlèn. The covariance matrices for which least squares is best linear unbiased. *Scand. J. Statist.*, 2:85–90, 1975.

[39] Q. Paris. The dual of the least-squares method. *Open J. Statist.*, 5:658–664, 2015.

[40] R. Penrose. A generalized inverse for matrices. *Proc. Cambridge Phil. Soc.*, 51:406–413, 1955.

[41] S. Puntanen and G.P.H. Styan. The equality of the ordinary least squares estimator and the best linear unbiased estimator, with comments by O. Kempthorne and S.R. Searle, and a reply by the authors. *Amer. Statistican*, 43:153–164, 1989.

[42] S. Puntanen, G.P.H. Styan, and J. Isotalo. *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty.* Springer, Berlin, 2011.

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

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

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

[46] C.R. Rao. *Linear Statistical Inference and Its Applications.* Wiley, New York, 1973.

[47] S.R. Searle. Additional results concerning estimable functions and generalized inverse matrices. *J. Roy. Statist. Soc. Ser. B*, 27:486–490, 1965.

[48] S.R. Searle. *Linear Models.* Wiley, New York, 1971.

[49] S.R. Searle. The infusion of matrices into statistics. *Bull. Internat. Lin. Alg. Soc.*, 24:25–32, 2000.

[50] J. Seely. Linear spaces and unbiased estimation. *Ann. Math. Statist.*, 41:1725–1734, 1970.

[51] J. Seely. Estimability and linear hypotheses. *Amer. Statistican*, 31:121–123, 1977.

[52] J. Seely and D. Birkes. Estimability in partitioned linear models. *Ann. Statist.*, 8:399–406, 1980.

[53] I. Stewart and H.P. Wynn. The estimability structure of linear models and submodels. *J. Roy. Stat. Soc. Ser. B*, 43:197–207, 1981.

[54] S.M. Stigler. Gauss and the invention of least squares. *Ann. Statist.*, 9:465–474, 1981.

[55] Y. Tian. Using rank formulas to characterize equalities for Moore–Penrose inverses of matrix products. *Appl. Math. Comput.*, 147:581–600, 2004.

[56] Y. Tian. Some decompositions of OLSEs and BLUEs under a partitioned linear model. *Internat. Statist. Rev.*, 75:224–248, 2007.

[57] Y. Tian. On an additive decomposition of the BLUE in a multiple partitioned linear model. *J. Multivariate Anal.*, 100:767–776, 2009.

[58] Y. Tian. On equalities of estimations of parametric functions under a general linear model and its restricted models. *Metrika*, 72:313–330, 2010.

[59] Y. Tian. Characterizing relationships between estimations under a general linear model with explicit and implicit restrictions by rank of matrix. *Comm. Statist. Theory Methods*, 41:2588–2601, 2012.

[60] Y. Tian. On properties of BLUEs under general linear regression models. *J. Statist. Plann. Inference*, 143:771–782, 2013.

[61] Y. Tian, M. Beisiegel, E. Dagenais, and C. Haines. On the natural restrictions in the singular Gauss–Markov model. *Stat. Papers*, 49:553–564, 2008.

[62] Y. Tian and B. Jiang. Equalities for estimators of partial parameters under linear model with restrictions. *J. Multivariate Anal.*, 143:299–313, 2016.

[63] Y. Tian and D.P. Wiens. The equalities of ordinary least-squares estimators and best linear unbiased estimators for the restricted linear model. *Stat. Prob. Lett.*, 76:1165–1272, 2006.

[64] Y. Tian and J. Zhang. Some equalities for estimations of partial coefficients under a general linear regression model. *Stat. Papers*, 52:911–920, 2011.

[65] Y. Tian and X. Zhang. On connections among OLSEs and BLUEs of whole and partial parameters under a general linear model. *Stat. Prob. Lett.*, 112:105–112, 2016.

[66] D.M. Young, P.L. Odell, and W. Hahn. Nonnegative-defininite covariance structures for which the blu, wls, and ls estimators are equal. *Stat. Prob. Lett.*, 49:271–276, 2000.