Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 8(7): 15189–15200.
DOI: 10.3934/math.2023775
Received: 22 February 2023
Revised: 11 March 2023
Accepted: 31 March 2023
Published: 24 April 2023

Research article

An effective treatment of adding-up restrictions in the inference of a general linear model

Yongge Tian*

College of Business and Economics, Shanghai Business School, Shanghai 201400, China

* **Correspondence:** Email: yongge.tian@gmail.com; Tel: +862131790567.

Abstract: This article offers a general procedure of carrying out estimation and inference under a linear statistical model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ with an adding-up restriction $\mathbf{Ay} = \mathbf{b}$ to the observed random vector $\mathbf{y}$. We first propose an available way of converting the adding-up restrictions to a linear matrix equation for $\boldsymbol{\beta}$ and a matrix equality for the covariance matrix of the error term $\boldsymbol{\varepsilon}$, which can help in combining the two model equations in certain consistent form. We then give the derivations and presentations of analytic expressions of the ordinary least-squares estimator (OLSE) and the best linear unbiased estimator (BLUE) of parametric vector $\mathbf{K}\boldsymbol{\beta}$ using various analytical algebraic operations of the given vectors and matrices in the model.

Keywords: adding-up restrictions; BLUE; estimability; general linear model; OLSE
Mathematics Subject Classification: 62H12, 62H20, 62J05

1. Introduction

Throughout, let $\mathbb{R}^{m\times n}$ stand for the collection of all $m \times n$ matrices over the field of real numbers; $\mathbf{A}'$, $r(\mathbf{A})$, 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; and let $\mathbf{I}_m$ denote the identity matrix of order m. Given an $\mathbf{A} \in \mathbb{R}^{m\times n}$, the Moore–Penrose generalized inverse of $\mathbf{A}$, 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, let $\mathbf{P_A}$, $\mathbf{E_A}$, and $\mathbf{F_A}$ stand for the three orthogonal projectors (symmetric idempotent matrices) $\mathbf{P_A} = \mathbf{AA}^+$, $\mathbf{E_A} = \mathbf{A}^\perp = \mathbf{I}_m - \mathbf{AA}^+$, and $\mathbf{F_A} = \mathbf{I}_n - \mathbf{A}^+\mathbf{A}$, which will help in briefly denoting calculation processes related to generalized inverses of matrices. Further information about the orthogonal projectors $\mathbf{P_A}$, $\mathbf{E_A}$, and $\mathbf{F_A}$ with their applications in the linear statistical models can be found, e.g., in [5, 9]. 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. For more results on the Löwner partial ordering of real symmetric matrices and its applications in statistical analysis, see e.g., [9, 17].

Consider a general linear model

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

where it is assumed that $\mathbf{y} \in \mathbb{R}^{n\times 1}$ is a vector of observable random variables, $\mathbf{X} \in \mathbb{R}^{n\times p}$ is a known model matrix of arbitrary rank ($0 \leq r(\mathbf{X}) \leq \min\{n, p\}$), $\boldsymbol{\beta} \in \mathbb{R}^{n\times 1}$ is a vector of fixed but unknown parameters, σ^2 is an arbitrary positive scaling factor, and $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ is a known nonnegative definite matrix of arbitrary rank ($0 \leq r(\boldsymbol{\Sigma}) \leq n$), for example $\boldsymbol{\Sigma} = \mathbf{I}$.

Below, we present some background details of the work. For a variety of reasons, statisticians may meet with the situation where certain restrictions are imposed on the unknown coefficient vector $\boldsymbol{\beta}$ and the observed random vector $\mathbf{y}$ in (1.1). For example, it is a regular case to add a system of linear matrix equations $\mathbf{B}\boldsymbol{\beta} = \mathbf{c}$ to the unknown parameter vector under the assumption in (1.1), and there have been plenty of approaches and discussions addressing how to carry out statistical inference under linear models with restrictions to their unknown parameters. In addition to the situations imposing restrictions on unknown parameters, it is necessary to take into account as well the situations of adding certain limitations and restrictions upon observed random variables in the model from theoretical and applied points of view. Under the model assumption in (1.1), one of such considerations is assuming that the observed random vector $\mathbf{y}$ satisfies a consistent linear matrix equation

$$\mathbf{A}\mathbf{y} = \mathbf{b}, \tag{1.2}$$

where it is assumed that $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a known matrix with $\mathrm{rank}(\mathbf{A}) = k \leq \min\{m, n\}$ and $\mathbf{b} \in \mathbb{R}^{m\times 1}$ is a known vector with $\mathbf{b} \in \mathscr{R}(\mathbf{A})$. A matrix equation as given in (1.2) is usually called adding-up restrictions to $\mathbf{y}$ in (1.1) in the literature. Clearly, the adding-up restrictions include $\mathbf{w}'\mathbf{y} = a$ or 1 etc. as its special cases, where $\mathbf{w}$ is a column vector. This kind of plausible-sounding restrictions do exist in statistical practice, which were noticed and approached in certain fields of applied statistics and attracted sort of consideration. For example, economists explored some situations of the specific kind where certain adding-up restrictions appeared, in which they presented fitting descriptions of how to think the restrictions, and gave some of their solutions to a number of corresponding estimation and inference problems. We refer the reader to [3,4,10,13] for more information on the appearance of such kind of adding-up restrictions. However, the general situation depicted in (1.2) has not been properly approached in the statistical literature, yet the very process of mathematical and statistical approaches of the adding-up restrictions remains hidden.

For the purpose of making inference in the contexts of (1.1) and (1.2), we merge the two model equations in the following form

$$\mathscr{N}:\ \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \ \ \mathbf{A}\mathbf{y} = \mathbf{b}, \ \ \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \ \ \mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{\Sigma}. \tag{1.3}$$

In this case, we can substitute (1.1) into (1.2) to lead to $\mathbf{A}\mathbf{X}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\varepsilon} = \mathbf{b}$, and rewrite (1.3) in the following equivalent form

$$\mathscr{N}:\ \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \ \ \mathbf{b} = \mathbf{A}\mathbf{X}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\varepsilon}, \ \ \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \ \ \mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\boldsymbol{\Sigma}. \tag{1.4}$$

Given the model equations as such in (1.3) or (1.4), we are confronted with the task of how to properly merge the adding-up restrictions in the estimation and inference procedure of the unknown parameter vector $\boldsymbol{\beta}$. There have been several attempts in the past to solve the problem of choosing possible merge procedures. Unfortunately, when faced to such a question, professional statisticians have no definitive answers or rather they have different methods which have been suggested and implemented in the literature, none of them being completely convincing in the sense that it has been shown to be better than the others. The purpose of this article is to focus attention on dealing with the adding-up restrictions in (1.3) with some new thoughts and methodologies. The author will offer a feasible algebraic method to reconcile the second adding-up restrictions with the first regression equation in (1.3), and then use the method to solve some basic estimation and inference problems associated with $\mathscr{N}$.

The rest of this paper is organized as follows. In Section 2, we introduce some basic formulas, facts, and results in matrix theory, as well as two groups of existing results related to the ordinary least-squares estimators (OLSE) and the best linear unbiased estimators (BLUEs) of unknown parametric vectors under (1.1). In Section 3, we show how to transform $\mathscr{N}$ in (1.3) into two kinds of linear models with implicit and explicit restrictions to the unknown parameter vector $\boldsymbol{\beta}$ respectively via certain suitable equivalent explanation of the adding-up restrictions. In Sections 4 and 5, we presents the description of the estimability of unknown parametric vector $\mathbf{K}\boldsymbol{\beta}$ under the transformed models, and give the definitions and the derivations of analytical expressions of the OLSEs and BLUEs of $\mathbf{K}\boldsymbol{\beta}$ through the transformed models. Section 6 gives some concluding remarks and a group of research problems concerning general linear models with adding-up restrictions.

2. Some preliminaries

In this section, we introduce some fundamental formulas and facts about matrix operations that have related applications to statistics, especially linear statistical models. It was properly known that the theory of generalized inverses of matrices is a major and dependable source of methods and techniques that was brought into the theory of linear statistical models for regression in 1950s, and thereby it played a key role for carrying out statistical estimation and inference in a wide variety of situations; see e.g., [2,9,14]. In this section, we shall present a group of well-known formulas, facts, and results in linear algebra and matrix theory, which we shall use as resource to simplify various matrix expressions that involve generalized inverses of matrices.

Lemma 2.1 ([7]). *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{m\times k}$, *and* $\mathbf{C} \in \mathbb{R}^{l\times n}$. *Then*,

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

(b) $r\begin{bmatrix}\mathbf{A}\\ \mathbf{C}\end{bmatrix} = r(\mathbf{A}) + r(\mathbf{C F_A}) = r(\mathbf{C}) + r(\mathbf{A F_C})$.

In particular,

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

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

Lemma 2.2 ([8]). *Let* $\mathbf{A} \in \mathbb{R}^{m\times n}$ *and* $\mathbf{B} \in \mathbb{R}^{m\times k}$, *Then*, *the linear matrix equation* $\mathbf{AX} = \mathbf{B}$ *is solvable for* $\mathbf{X}$ *if and only if* $r[\mathbf{A}, \mathbf{B}] = r(\mathbf{A})$, *or equivalently*, $\mathbf{AA^+B} = \mathbf{B}$. *In this case*, *the general solution of the equation can be written in the parametric form* $\mathbf{X} = \mathbf{A^+B} + (\mathbf{I}_n - \mathbf{A^+A})\mathbf{U}$, *where* $\mathbf{U} \in \mathbb{R}^{n\times k}$ *is an arbitrary matrix*.

There were different established inference theories and methods that we can adopt to estimate the unknown parameter vector $\boldsymbol{\beta}$ in (1.1), the two best-known tools were OLSEs and BLUEs. We turn now to reviewing some basic definitions and existing facts in linear model theory regarding the estimability, as well as the OLSEs and BLUEs of a given unknown parametric vector under (1.1), see e.g., [1, 9, 11, 12, 18].

Definition 2.3. Let $\mathscr{M}$ be as given in (1.1) and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given. The vector $\mathbf{K}\boldsymbol{\beta}$ of parametric functions is said to be *estimable* under $\mathscr{M}$ if there exists an $\mathbf{L} \in \mathbb{R}^{k\times n}$ such that $\mathrm{E}(\mathbf{Ly} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0}$ holds for all $\boldsymbol{\beta}$ in $\mathscr{M}$.

Definition 2.4. Let $\mathscr{M}$ be as given in (1.1), and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given.

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

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathrm{argmin}}(\,\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\,)'(\,\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\,). \tag{2.1}$$

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

(b) The BLUE of the vector of parametric functions $\mathbf{K}\boldsymbol{\beta}$ under (1.1), denoted by $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$, is defined to be linear statistic $\mathbf{Ly}$, where $\mathbf{L}$ is a matrix such that $\mathrm{Cov}(\mathbf{Ly} - \mathbf{K}\boldsymbol{\beta}) = \min$ in the Löwner partial ordering subject to $\mathrm{E}(\mathbf{Ly} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0}$.

As we know that the concepts of OLSE and BLUE have a long history and deep roots in parametric regression analysis, both of which have many nice and optimal algebraic and statistical properties, and therefore are the most welcome linear statistical inference techniques in parametric regression theory and the related applications. The conventionality of OLSEs/BLUEs under linear regression models really attracted statisticians' attention in the development of regression theory, and numerous formulas and facts regarding the OLSEs/BLUEs of $\boldsymbol{\beta}$ and $\mathbf{K}\boldsymbol{\beta}$ under (1.1) were established via various precise and analytical algebraic operations of the given vectors and matrices and their generalized inverses. Specifically, the results in the following two lemmas were highly recognized in the domain of linear statistical models.

Lemma 2.5. *Let $\mathscr{M}$ be as given in* (1.1), *and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given. Then, the general expression of OLSEs of $\boldsymbol{\beta}$ in $\mathscr{M}$ can be written as*

$$\mathrm{OLSE}_{\mathscr{M}}(\boldsymbol{\beta}) = \mathbf{X}^{+}\mathbf{y} + \mathbf{F}_{\mathbf{X}}\mathbf{v}, \tag{2.2}$$

where $\mathbf{v} \in \mathbb{R}^{p\times 1}$ is arbitrary; *and the OLSE of $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{M}$ can be written as*

$$\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathbf{X}^{+}\mathbf{y} + \mathbf{K}\mathbf{F}_{\mathbf{X}}\mathbf{v}. \tag{2.3}$$

Lemma 2.6. *Let $\mathscr{M}$ be as given in* (1.1), $\mathbf{K} \in \mathbb{R}^{k\times p}$, *and suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{M}$, namely, $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$. Then, the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{M}$ can be written as*

$$\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{K};\mathbf{X};\boldsymbol{\Sigma}}\mathbf{y}, \tag{2.4}$$

where $\mathbf{P}_{\mathbf{K};\mathbf{X};\boldsymbol{\Sigma}}$ is the solution of the matrix equation

$$\mathbf{G}[\,\mathbf{X},\ \boldsymbol{\Sigma}\mathbf{X}^{\perp}\,] = [\,\mathbf{K},\ \mathbf{0}\,]. \tag{2.5}$$

This equation is always solvable for $\mathbf{G}$, *that is*, $\mathscr{R}([\,\mathbf{K},\ \mathbf{0}\,]') \subseteq \mathscr{R}([\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}]')$. *In this case*, *the general solution of* (2.5) *can be expressed as*

$$\mathbf{P}_{\mathbf{K};\mathbf{X};\mathbf{\Sigma}} = [\mathbf{K},\ \mathbf{0}][\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}]^{+} + \mathbf{U}[\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}]^{\perp}, \tag{2.6}$$

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

(a) $r[\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}] = r[\mathbf{X},\ \mathbf{\Sigma}]$ *and* $\mathscr{R}[\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}] = \mathscr{R}[\mathbf{X},\ \mathbf{\Sigma}]$.
(b) *The product* $\mathbf{P}_{\mathbf{K};\mathbf{X};\mathbf{\Sigma}}\mathbf{\Sigma}$ *can be uniquely written as* $\mathbf{P}_{\mathbf{K};\mathbf{X};\mathbf{\Sigma}}\mathbf{\Sigma} = [\mathbf{K},\ \mathbf{0}][\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}]^{+}\mathbf{\Sigma}$.
(c) *The expectation and covariance matrix of* $\mathrm{BLUE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta})$ *are given by*

$$\mathrm{E}(\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})) = \mathbf{K}\boldsymbol{\beta} \ \ \textit{and} \ \ \mathrm{Cov}(\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})) = [\mathbf{K},\ \mathbf{0}][\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}]^{+}\mathbf{\Sigma}([\mathbf{K},\ \mathbf{0}][\mathbf{X},\ \mathbf{\Sigma}\mathbf{X}^{\perp}]^{+})'.$$

(d) *The matrix* $\mathbf{P}_{\mathbf{K};\mathbf{X};\mathbf{\Sigma}}$ *is unique if and only if* $r[\mathbf{X},\ \mathbf{\Sigma}] = n$.
(e) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta})$ *is unique if and only if if and only if* $\mathbf{y} \in \mathscr{R}[\mathbf{X},\ \mathbf{\Sigma}]$ *holds with probability* 1.

Apparently, the representations and derivations of the OLSEs and BLUEs in the above two existing lemmas do not require to specify the mathematical forms of the distributions of the error term vector $\boldsymbol{\varepsilon}$, and thereby we are able to actively utilize the clear and exact expressions of the OLSEs and BLUEs to solve many algebraic and computational problems in the context of (1.1).

3. Transformations of $\mathscr{N}$ into linear models with implicit and explicit restrictions to unknown parameters

Because the adding-up equation is directly put on the random vector $\mathbf{y}$ in (1.1), where no information of $\boldsymbol{\beta}$ explicitly appears in the equation, we should be careful to include the equation in the estimation and inference procedure of (1.1). In other words, we have to seek some alternative methods to approach estimation and inference problems of unknown parameters in the model. To this purpose, we show in this section how to make use of the complete information associated with the adding-up restrictions and to convert (1.3) into certain ordinary linear models with implicit and explicit restrictions to the unknown parametric vector $\boldsymbol{\beta}$, respectively.

Given that $\mathbf{y}$ in (1.1) is a random vector, we then take the expectation and covariance matrix of both sides of the equation $\mathbf{A}\mathbf{y} - \mathbf{b} = \mathbf{0}$ with respect to $\mathbf{y}$ to obtain

$$\mathrm{E}(\mathbf{A}\mathbf{y} - \mathbf{b}) = \mathbf{A}\mathbf{X}\boldsymbol{\beta} - \mathbf{b} = \mathbf{0} \ \ \text{and} \ \ \mathrm{Cov}(\mathbf{A}\mathbf{y} - \mathbf{b}) = \sigma^2\mathbf{A}\mathbf{\Sigma}\mathbf{A}' = \mathbf{0}. \tag{3.1}$$

Since the matrix $\mathbf{\Sigma}$ in (1.3) is positive semi-definite, it is easy to verify that the matrix equality $\mathbf{A}\mathbf{\Sigma}\mathbf{A}' = \mathbf{0}$ is equivalent to $\mathbf{\Sigma} = \mathbf{F}_{\mathbf{A}}\mathbf{\Sigma}\mathbf{F}_{\mathbf{A}}$. The adding-up equation in (1.2) thereby suggests the following facts

$$\mathbf{A}\mathbf{X}\boldsymbol{\beta} = \mathbf{b} \ \text{and} \ \mathbf{\Sigma} = \mathbf{F}_{\mathbf{A}}\mathbf{\Sigma}\mathbf{F}_{\mathbf{A}}. \tag{3.2}$$

This treatment is by no means profound and difficult to understand under the assumption in (1.3), and thus we can view them as the best mathematical interpretation that can be given about the adding-up restrictions. Recognizing the key role of (3.2) in the interpretation of (1.2), we are able subsequently to deal with the adding-up equation in carrying out inference under (1.3). There are basically two algebraic methods to merge the adding-up equation into (1.1) via (3.2). Below, we perspicuously illustrate the algebraic processes.

(I) Firstly, substituting the first equation into the second equation in (1.3) and noting (3.2), we can equivalently rewrite (1.3) in the following implicitly restricted linear model

$$\mathscr{N}_a: \begin{bmatrix}\mathbf{y}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{X}\\ \mathbf{AX}\end{bmatrix}\boldsymbol{\beta} + \begin{bmatrix}\boldsymbol{\varepsilon}\\ \mathbf{A}\boldsymbol{\varepsilon}\end{bmatrix}, \ \mathrm{E}\begin{bmatrix}\boldsymbol{\varepsilon}\\ \mathbf{A}\boldsymbol{\varepsilon}\end{bmatrix} = \mathbf{0}, \ \mathrm{Cov}\begin{bmatrix}\boldsymbol{\varepsilon}\\ \mathbf{A}\boldsymbol{\varepsilon}\end{bmatrix} = \sigma^2\begin{bmatrix}\mathbf{F_A\Sigma F_A} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}. \tag{3.3}$$

(II) Also replacing $\mathbf{Ay} = \mathbf{b}$ and $\boldsymbol{\Sigma}$ in (1.3) with (3.2) produces with probability 1 the following explicitly restricted linear model

$$\mathscr{N}_b: \ \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \ \ \mathbf{AX}\boldsymbol{\beta} = \mathbf{b}, \ \ \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \ \ \mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{F_A\Sigma F_A}. \tag{3.4}$$

Apparently, the two alternative forms in (3.3) and (3.4) are in compliance with $\mathscr{N}$ in (1.3), and thereby they can conveniently help solve various estimation and inference problems under $\mathscr{N}$. It is easily seen that (3.3) and (3.4) are nothing but two ordinary linear statistical models with implicit and explicit restrictions to unknown parameters in the models. This fact enables us to adopt different approach processes to carry out common estimation and inference under (1.3) via (3.3) and (3.4), and thereby this alternative way in fact makes clear insights actionable into the connotation of the model in (1.3).

As a classic subject of study in regression analysis, there has been some general discussion regarding estimation and inference problems of a linear statistical model with implicit and explicit restrictions to unknown parameters in the model; see, e.g., [19] and references therein. In light of the existing theory pertaining to this topic, we are now able to make statistical inference of (1.3) via the two alternative forms in (3.3) and (3.4) through the well-organized employment of ordinary theory and methodology of dealing linear regression models under various assumptions.

4. Estimation results under $\mathscr{N}_a$

For convenience of representation, we adopt the notation

$$\widehat{\mathbf{y}} = \begin{bmatrix}\mathbf{y}\\ \mathbf{b}\end{bmatrix}, \ \ \widehat{\mathbf{X}} = \begin{bmatrix}\mathbf{X}\\ \mathbf{AX}\end{bmatrix}, \ \ \widehat{\boldsymbol{\Sigma}} = \begin{bmatrix}\mathbf{F_A\Sigma F_A} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}$$

in the sequel. We first describe the consistency problem in the context of (3.3). Note that the matrix equality

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

holds from the definition of the Moore–Penrose generalized inverse. Therefore, it turns out under the assumptions in (3.2) that

$$\mathrm{E}([\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}]^{+}\widehat{\mathbf{y}} - \widehat{\mathbf{y}}) = [\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}]^{+}\widehat{\mathbf{X}}\boldsymbol{\beta} - \widehat{\mathbf{X}}\boldsymbol{\beta} = \mathbf{0},$$
$$\mathrm{Cov}([\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}]^{+}\widehat{\mathbf{y}} - \widehat{\mathbf{y}}) = \sigma^2([\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}]^{+} - \mathbf{I})\widehat{\boldsymbol{\Sigma}}([\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}]^{+} - \mathbf{I})' = \mathbf{0}.$$

These two equalities imply $[\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}]^{+}\widehat{\mathbf{y}} = \widehat{\mathbf{y}}$ holds with probability 1, or equivalently,

$$\widehat{\mathbf{y}} \in \mathscr{R}[\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}}] \tag{4.1}$$

holds with probability 1. In view of this fact, we adopt the following definition.

Definition 4.1. $\mathscr{N}_a$ in (3.3) is said to be *consistent* if (4.1) holds with probability 1.

The OLSEs and BLUEs of unknown parameters in a given linear statistical model were recognized as two principal estimations in the domain of linear regression models, which were deeply approached and utilized in the development of statistical science. In the following, we introduce the definitions of the OLSEs and BLUEs of vectors of parametric functions, and then presents the exact and analytical formulas for calculating the OLSEs and BLUEs under (3.3).

Definition 4.2. Let $\mathscr{N}_a$ be as given in (3.3), and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given. The vector $\mathbf{K}\boldsymbol{\beta}$ of parametric functions is said to be *estimable* under $\mathscr{N}_a$ if there exists an $\mathbf{L} \in \mathbb{R}^{k\times(n+m)}$ such that $\mathrm{E}(\mathbf{L}\widehat{\mathbf{y}} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0}$ holds for all $\boldsymbol{\beta}$ under $\mathscr{N}_a$.

Definition 4.3. Let $\mathscr{N}_a$ be as given in (3.3), and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given.

(a) The OLSE of the parametric vector $\boldsymbol{\beta}$ under (3.3), denoted by $\mathrm{OLSE}_{\mathscr{N}_a}(\boldsymbol{\beta})$, is defined to be

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\mathrm{argmin}}(\widehat{\mathbf{y}} - \widehat{\mathbf{X}}\boldsymbol{\beta})'(\widehat{\mathbf{y}} - \widehat{\mathbf{X}}\boldsymbol{\beta}).$$

The OLSE of $\mathbf{K}\boldsymbol{\beta}$ under (3.3) is defined to be $\mathrm{OLSE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\mathrm{OLSE}_{\mathscr{N}_a}(\boldsymbol{\beta})$.

(b) The BLUE of the vector of parametric functions $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{N}$, denoted by $\mathrm{BLUE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})$, is defined to be a linear statistic $\mathbf{L}\widehat{\mathbf{y}}$, where $\mathbf{L}$ is a matrix such that $\mathrm{Cov}(\mathbf{L}\widehat{\mathbf{y}} - \mathbf{K}\boldsymbol{\beta}) = \min$ in the Löwner partial ordering subject to $\mathrm{E}(\mathbf{L}\widehat{\mathbf{y}} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0}$.

Now applying the above definitions to $\mathscr{N}_a$ in (3.3), we obtain the following results.

Theorem 4.4. *Let $\mathscr{N}_a$ be as given in* (3.3), *and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given. Then, $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{N}_a$ $\Leftrightarrow \mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$. In particular, $\mathbf{X}\boldsymbol{\beta}$ is always estimable under $\mathscr{N}_a$.*

Proof. It follows from $\mathrm{E}(\mathbf{L}\widehat{\mathbf{y}} - \mathbf{K}\boldsymbol{\beta}) = \mathbf{0} \Leftrightarrow \mathbf{L}\widehat{\mathbf{X}}\boldsymbol{\beta} - \mathbf{K}\boldsymbol{\beta} = \mathbf{0}$ for all $\boldsymbol{\beta} \Leftrightarrow \mathbf{L}\widehat{\mathbf{X}} = \mathbf{K} \Leftrightarrow \mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\widehat{\mathbf{X}}') \Leftrightarrow \mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$ by Lemma 2.2. □

Referring to Lemmas 2.5 and 2.6, we obtain the following two results about the OLSEs and BLUEs under (3.3).

Theorem 4.5. *Let $\mathscr{N}_a$ be as given in* (3.3) *and suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{N}_a$. Then, the OLSE of $\boldsymbol{\beta}$ under $\mathscr{N}_a$ can be written as $\mathrm{OLSE}_{\mathscr{N}_a}(\boldsymbol{\beta}) = \widehat{\mathbf{X}}^{+}\widehat{\mathbf{y}} + \mathbf{F}_{\mathbf{X}}\mathbf{v}$, where $\mathbf{v} \in \mathbb{R}^{p\times 1}$ is arbitrary; and the OLSE of $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{N}_a$ can be uniquely written as*

$$\mathrm{OLSE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\widehat{\mathbf{X}}^{+}\widehat{\mathbf{y}},\ \ \mathrm{E}(\mathrm{OLSE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})) = \mathbf{K}\boldsymbol{\beta},\ \ \mathrm{Cov}(\mathrm{OLSE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})) = \sigma^2\mathbf{K}\widehat{\mathbf{X}}^{+}\widehat{\boldsymbol{\Sigma}}(\mathbf{K}\widehat{\mathbf{X}}^{+})'.$$

Theorem 4.6. *Let $\mathscr{N}_a$ be as given in* (3.3) *and suppose $\mathbf{K}\boldsymbol{\beta}$ is estimable under $\mathscr{N}_a$. Then, the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under $\mathscr{N}_a$ can be written as*

$$\mathrm{BLUE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{P}_{\mathbf{K};\widehat{\mathbf{X}};\widehat{\boldsymbol{\Sigma}}}\widehat{\mathbf{y}},$$

where $\mathbf{P}_{\mathbf{K};\widehat{\mathbf{X}};\widehat{\boldsymbol{\Sigma}}}$ is the solution of the matrix equation $\mathbf{G}[\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}\mathbf{X}}^{\perp}] = [\mathbf{K}, \mathbf{0}]$. This equation is always solvable for $\mathbf{G}$, that is, $\mathscr{R}([\mathbf{K}, \mathbf{0}]') \subseteq \mathscr{R}([\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}\mathbf{X}}^{\perp}]')$. In this case, the general solution of the matrix equation can be expressed as

$$\mathbf{G} = \mathbf{P}_{\mathbf{K};\widehat{\mathbf{X}};\widehat{\boldsymbol{\Sigma}}} = [\mathbf{K}, \mathbf{0}][\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}\mathbf{X}}^{\perp}]^{+} + \mathbf{U}\mathbf{E}_{[\widehat{\mathbf{X}}, \widehat{\boldsymbol{\Sigma}\mathbf{X}}^{\perp}]},$$

where $\mathbf{U} \in \mathbb{R}^{k\times(n+m)}$ is arbitrary. Moreover, the following results hold.

(a) $r[\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma X}^{\perp}}] = r[\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma}}]$ *and* $\mathscr{R}[\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma X}^{\perp}}] = \mathscr{R}[\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma}}]$.
(b) *The product* $\mathbf{P}_{\mathbf{K};\widehat{\mathbf{X}};\widehat{\mathbf{\Sigma}}}\widehat{\mathbf{\Sigma}}$ *can be uniquely written as* $\mathbf{P}_{\mathbf{K};\widehat{\mathbf{X}};\widehat{\mathbf{\Sigma}}}\widehat{\mathbf{\Sigma}} = [\mathbf{K}, \mathbf{0}][\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma X}^{\perp}}]^{+}\widehat{\mathbf{\Sigma}}$.
(c) *The expectation and covariance matrix of* $\mathrm{BLUE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})$ *are given by*

$$\mathrm{E}(\mathrm{BLUE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})) = \mathbf{K}\boldsymbol{\beta} \ \ \textit{and} \ \ \mathrm{Cov}(\mathrm{BLUE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})) = [\mathbf{K}, \mathbf{0}][\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma X}^{\perp}}]^{+}\widehat{\mathbf{\Sigma}}([\mathbf{K}, \mathbf{0}][\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma X}^{\perp}}]^{+})'.$$

(d) *The matrix* $\mathbf{P}_{\mathbf{K};\widehat{\mathbf{X}};\widehat{\mathbf{\Sigma}}}$ *is unique if and only if* $r[\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma}}] = m + n$.
(e) $\mathrm{BLUE}_{\mathscr{N}_a}(\mathbf{K}\boldsymbol{\beta})$ *is unique if and only if* $\widehat{\mathbf{y}} \in \mathscr{R}[\widehat{\mathbf{X}}, \widehat{\mathbf{\Sigma}}]$ *holds with probability* 1.

5. Estimation results under $\mathscr{N}_b$

In what follows, we denote $\widetilde{\mathbf{A}} = \mathbf{AX}$ and $\widetilde{\mathbf{\Sigma}} = \mathbf{F_A \Sigma F_A}$. Recall a well-known fact that the matrix equation $\widetilde{\mathbf{A}}\boldsymbol{\beta} = \mathbf{b}$ is solvable for $\boldsymbol{\beta}$ if and only if $\mathbf{b} \in \mathscr{R}(\widetilde{\mathbf{A}})$. By Lemma 2.2, the general solution of $\boldsymbol{\beta}$ and the corresponding $\mathbf{K}\boldsymbol{\beta}$ can be written in the following parametric forms

$$\boldsymbol{\beta} = \widetilde{\mathbf{A}}^{+}\mathbf{b} + \mathbf{F}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma}, \tag{5.1}$$

$$\mathbf{K}\boldsymbol{\beta} = \mathbf{K}\widetilde{\mathbf{A}}^{+}\mathbf{b} + \mathbf{K}\mathbf{F}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma}, \tag{5.2}$$

where $\boldsymbol{\gamma} \in \mathbb{R}^{p\times 1}$ is arbitrary. Substitution of (5.1) into (3.4) yields

$$\widetilde{\mathscr{N}_b}: \ \mathbf{z} = \mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \ \ \mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \ \ \mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2\widetilde{\mathbf{\Sigma}}, \tag{5.3}$$

where $\mathbf{z} = \mathbf{y} - \mathbf{X}\widetilde{\mathbf{A}}^{+}\mathbf{b}$. This is a new linear model with the unknown parameter vector $\boldsymbol{\gamma}$. Hence, the estimablility, the OLSE, and the BLUE of the vector of parametric functions $\mathbf{K}\mathbf{F}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma}$ can be obtained from various existing results as follows.

Definition 5.1. Let $\mathscr{N}_b$ be as given in (3.4), and let $\mathbf{K} \in \mathbb{R}^{k\times p}$ be given. The vector $\mathbf{K}\boldsymbol{\beta}$ of parametric functions is said to be *estimable* under $\mathscr{N}_b$ if there exist $\mathbf{L} \in \mathbb{R}^{k\times n}$ and $\mathbf{c} \in \mathbb{R}^{k\times 1}$ such that $\mathrm{E}(\mathbf{Ly}+\mathbf{c}-\mathbf{K}\boldsymbol{\beta}) = \mathbf{0}$ holds under $\mathscr{N}_b$.

Lemma 5.2. *Let* $\mathscr{N}_b$ *be as given in* (3.4), *and let* $\mathbf{K} \in \mathbb{R}^{k\times p}$ *be given. Then,* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* $\mathscr{N}_b$ *if and only if* $\mathscr{R}(\mathbf{K}') \subseteq \mathscr{R}(\mathbf{X}')$.

Theorem 5.3. *Let* $\mathscr{N}_b$ *be as given in* (3.4), *let* $\mathbf{K} \in \mathbb{R}^{k\times p}$ *be given, and suppose* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (3.4). *Then, the OLSE of* $\boldsymbol{\beta}$ *under* $\mathscr{N}_b$ *can be written as*

$$\mathrm{OLSE}_{\mathscr{N}_b}(\boldsymbol{\beta}) = (\widetilde{\mathbf{A}}^{+} - \mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{X}\widetilde{\mathbf{A}}^{+})\mathbf{b} + \mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{y} + \mathbf{F}_{\widetilde{\mathbf{A}}}\mathbf{F}_{\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}}}\mathbf{u}, \tag{5.4}$$

where $\mathbf{u} \in \mathbb{R}^{p\times 1}$ *is arbitrary. The OLSE of* $\mathbf{K}\boldsymbol{\beta}$ *under* $\mathscr{N}_b$ *can be uniquely written as*

$$\mathrm{OLSE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta}) = (\mathbf{K}\widetilde{\mathbf{A}}^{+} - \mathbf{K}\mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{X}\widetilde{\mathbf{A}}^{+})\mathbf{b} + \mathbf{K}\mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{y}, \tag{5.5}$$

$$\mathrm{E}(\,\mathrm{OLSE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})\,) = \mathbf{K}\boldsymbol{\beta}, \ \ \mathrm{Cov}(\,\mathrm{OLSE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})\,) = \sigma^2\mathbf{K}\mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\widetilde{\mathbf{\Sigma}}(\mathbf{K}\mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+})'.$$

Proof. According to Lemma 2.5, the OLSE of $\boldsymbol{\gamma}$ in (5.3) can be written as

$$\widehat{\boldsymbol{\gamma}} = (\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{z} + \mathbf{F}_{\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}}}\mathbf{u},$$

where $\mathbf{u} \in \mathbb{R}^{p\times 1}$ is arbitrary. Substitution of this formula into (5.1) gives the OLSE of $\boldsymbol{\beta}$ in (3.4):

$$\mathrm{OLSE}_{\mathscr{N}_b}(\boldsymbol{\beta}) = \widetilde{\mathbf{A}}^{+}\mathbf{b} + \mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{z} + \mathbf{F}_{\widetilde{\mathbf{A}}}\mathbf{F}_{\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}}}\mathbf{u} = (\widetilde{\mathbf{A}}^{+} - \mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{X}\widetilde{\mathbf{A}}^{+})\mathbf{b} + \mathbf{F}_{\widetilde{\mathbf{A}}}(\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}})^{+}\mathbf{y} + \mathbf{F}_{\widetilde{\mathbf{A}}}\mathbf{F}_{\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}}}\mathbf{u},$$

establishing (5.4) and (5.5). □

Theorem 5.4. *Let* $\mathscr{N}_b$ *be as given in* (3.4) *and suppose* $\mathbf{K}\boldsymbol{\beta}$ *is estimable under* (3.4). *Then, the BLUE of* $\mathbf{K}\boldsymbol{\beta}$ *under* $\mathscr{N}_b$ *can be written as*

$$\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta}) = (\mathbf{K} - \mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}\mathbf{X})\widetilde{\mathbf{A}}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}\mathbf{y}, \tag{5.6}$$

where $\mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}} = [\,\mathbf{KF}_{\widetilde{\mathbf{A}}},\ \mathbf{0}\,][\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}\,]^{+} + \mathbf{U}_1\mathbf{E}_{[\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\,\widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}\,]}$, *and* $\mathbf{U}_1 \in \mathbb{R}^{k\times n}$ *is arbitrary. Moreover, the following results hold.*

(a) $r[\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}] = r[\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}]$ *and* $\mathscr{R}[\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}] = \mathscr{R}[\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}]$.

(b) *The product* $\mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}\widetilde{\mathbf{\Sigma}}$ *can be uniquely written as* $\mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}\widetilde{\mathbf{\Sigma}} = [\mathbf{KF}_{\widetilde{\mathbf{A}}},\ \mathbf{0}][\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}]^{+}\widetilde{\mathbf{\Sigma}}$.

(c) *The expectation and covariance matrix of* $\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})$ *are given by*

$$\mathrm{E}(\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})) = \mathbf{K}\boldsymbol{\beta} \ \ \textit{and} \ \ \mathrm{Cov}(\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})) = \mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}\widetilde{\mathbf{\Sigma}}\mathbf{P}'_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}.$$

(d) *The matrix* $\mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}$ *is unique if and only if* $r\begin{bmatrix}\mathbf{X} & \widetilde{\mathbf{\Sigma}}\\ \widetilde{\mathbf{A}} & \mathbf{0}\end{bmatrix} = r(\widetilde{\mathbf{A}}) + n$.

(e) $\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})$ *is unique if and only if* $\begin{bmatrix}\mathbf{y}\\ \mathbf{b}\end{bmatrix} \in \mathscr{R}\begin{bmatrix}\mathbf{X} & \widetilde{\mathbf{\Sigma}}\\ \widetilde{\mathbf{A}} & \mathbf{0}\end{bmatrix}$ *holds with probability* 1.

Proof. According to Lemma 2.6, the BLUE of $\mathbf{KF}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma}$ under (5.3) is given by

$$\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{KF}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma}) = \mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}\mathbf{z}.$$

Substitution of this BLUE into the equality in (5.2) gives the BLUE of $\mathbf{K}\boldsymbol{\beta}$ under (3.4)

$$\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta}) = \mathbf{K}\widetilde{\mathbf{A}}^{+}\mathbf{b} + \mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{KF}_{\widetilde{\mathbf{A}}}\boldsymbol{\gamma}) = \mathbf{K}\widetilde{\mathbf{A}}^{+}\mathbf{b} + \mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}(\,\mathbf{y} - \mathbf{X}\widetilde{\mathbf{A}}^{+}\mathbf{b}\,),$$

as required for (5.6).

Result (a) follows from Lemma 2.6(a). Result (b) follows from Lemma 2.6(b). Result (c) follows from (5.6).

It can be seen from (5.6) that $\mathbf{P}_{\mathbf{KF}_{\widetilde{\mathbf{A}}};\mathbf{XF}_{\widetilde{\mathbf{A}}};\widetilde{\mathbf{\Sigma}}}$ is unique if and only if $\mathbf{E}_{[\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\,\widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}\,]} = \mathbf{0}$, i.e., $r[\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}\,] = n$. Also see from (a) and Lemma 2.1(b) that $r[\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}\,] = r[\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\,] = r\begin{bmatrix}\mathbf{X} & \widetilde{\mathbf{\Sigma}}\\ \widetilde{\mathbf{A}} & \mathbf{0}\end{bmatrix} - r(\widetilde{\mathbf{A}})$, so that Result (d) follows.

It can be seen from (5.6) and that $\mathrm{BLUE}_{\mathscr{N}_b}(\mathbf{K}\boldsymbol{\beta})$ is unique if and only if $\mathbf{E}_{[\,\mathbf{XF}_{\widetilde{\mathbf{A}}},\,\widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}\,]}(\mathbf{y} - \mathbf{X}\widetilde{\mathbf{A}}^{+}\mathbf{b}) = \mathbf{0}$, i.e.,

$$r[\mathbf{y} - \mathbf{X}\widetilde{\mathbf{A}}^{+}\mathbf{b},\ \mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}] = r[\mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}] \tag{5.7}$$

holds with probability 1 by Lemma 2.1(c). In this situation, it is necessary to simplify the rank equality by removing the generalized inverses on both sides of (5.7). In fact by Lemma 2.1(b) and elementary block matrix operations,

$$r[\mathbf{y} - \mathbf{X}\widetilde{\mathbf{A}}^{+}\mathbf{b},\ \mathbf{XF}_{\widetilde{\mathbf{A}}},\ \widetilde{\mathbf{\Sigma}}\mathbf{E}_{\mathbf{XF}_{\widetilde{\mathbf{A}}}}] = r\begin{bmatrix}\mathbf{y} - \mathbf{X}\widetilde{\mathbf{A}}^{+}\mathbf{b} & \mathbf{X} & \widetilde{\mathbf{\Sigma}}\\ \mathbf{0} & \widetilde{\mathbf{A}} & \mathbf{0}\end{bmatrix} - r(\widetilde{\mathbf{A}}) = r\begin{bmatrix}\mathbf{y} & \mathbf{X} & \widetilde{\mathbf{\Sigma}}\\ \mathbf{b} & \widetilde{\mathbf{A}} & \mathbf{0}\end{bmatrix} - r(\widetilde{\mathbf{A}}),$$

$$r[\,\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}},\ \widetilde{\boldsymbol{\Sigma}}\,] = r[\,\mathbf{X}\mathbf{F}_{\widetilde{\mathbf{A}}},\ \widetilde{\boldsymbol{\Sigma}}\,] = r\begin{bmatrix} \mathbf{X} & \widetilde{\boldsymbol{\Sigma}} \\ \widetilde{\mathbf{A}} & \mathbf{0} \end{bmatrix} - r(\widetilde{\mathbf{A}}).$$

So that (5.7) is equivalent to $r\begin{bmatrix} \mathbf{y} & \mathbf{X} & \widetilde{\boldsymbol{\Sigma}} \\ \mathbf{b} & \widetilde{\mathbf{A}} & \mathbf{0} \end{bmatrix} = r\begin{bmatrix} \mathbf{X} & \widetilde{\boldsymbol{\Sigma}} \\ \widetilde{\mathbf{A}} & \mathbf{0} \end{bmatrix}$, i.e., $\begin{bmatrix} \mathbf{y} \\ \mathbf{b} \end{bmatrix} \in \mathscr{R}\begin{bmatrix} \mathbf{X} & \widetilde{\boldsymbol{\Sigma}} \\ \widetilde{\mathbf{A}} & \mathbf{0} \end{bmatrix}$ holds by Lemma 2.1(c), as required for Result (e). □

6. Conclusions

The author proposed and investigated some estimation and inference problems regarding a linear statistical model with adding-up restrictions to the observable random variables in the model, and obtained a group of formulas and facts about estimations and inferences in the context of (1.3), including the computational processes to derive the OLSEs and BLUEs under the model assumptions via a series of careful algebraic operations of the given vectors and matrices. Based on the findings obtained in the preceding sections, there is no doubt to say that this approach clearly demonstrates a normal procedure of dealing with adding-up restrictions in the context of linear statistical models. Hopefully, this study will add to what is out there with regard to the subject, and also is compatible with the previous contributions.

Given the resolutions to the adding-up restrictions, we may say some additional remarks pertaining to further research problems under the model assumption. Recall that the OLSEs and BLUEs are defined by different optimality criteria in mathematics and statistics. Therefore, their expressions and properties are not necessarily the same, and thereby it is natural to seek possible connections between these estimation results. It is, in fact, a subject area in regression analysis is to characterize relationships between different OLSEs and BLUEs, which has deep roots with strong statistical explanation and usefulness in the domain of linear statistical models and applications; see, e.g., [6, 15, 16, 20, 21] and references therein for the background and study of this subject. Based on the exact and analytical expressions of OLSEs and BLUEs obtained, we can consider in depth various additional topics in the statistical inference of general linear statistical models with adding-up restrictions. Particularly, it is natural to speculate on the relationship between the OLSEs and BLUEs under the two models in (1.1) and (1.3), and therefore we put forward the following five clear and reasonable equalities between the OLSEs and BLUEs under the two models in (1.1) and (1.3):

(a) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta})$,
(b) $\mathrm{OLSE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta})$,
(c) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{OLSE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta})$,
(d) $\mathrm{BLUE}_{\mathscr{M}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta})$,
(e) $\mathrm{OLSE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}_{\mathscr{N}}(\mathbf{K}\boldsymbol{\beta})$.

The five equalities describe the direct relevances of the four estimators under the two models, and thus it would be of interest to make a deep-going study of the equalities from theoretical and practical points of view. As a matter of fact, these kinds of equalities were properly considered with clear objective intention in the statistical inference of linear statistical models in the past several decades. Unquestionably, the equalities in (a)–(e) not only can be classified as certain statistical inference problems, but also can be alternatively converted to certain matrix equality problems in term of the clear and analytical expressions of the OLSEs and BLUEs obtained in the preceding sections. In order to resolve these proposed equivalence problems, we need to prepare a complex of preliminary methods and techniques in matrix algebra, including many formulas and facts about generalized inverses of matrices and the matrix rank methodology. Nevertheless, this sort of future investigations together with the current theoretical and methodological advances in this research area will definitely provide considerable insight into the intrinsic natures hidden behind the adding-up restrictions, so that we believe that the algebraic treatments presented in this article will sufficiently prompt other similar studies regarding different kinds of regression models with adding-up restrictions to observable random variables under various assumptions. The five equalities describe the direct relevances of the four estimators under the two models, and thus it would be of interest to make a deep-going study of the equalities from theoretical and practical points of view. As a matter of fact, these kinds of equalities were properly considered with clear objective intention in the statistical inference of linear statistical models in the past several decades. Unquestionably, the equalities in (a)–(e) not only can be classified as certain statistical inference problems, but also can be alternatively converted to certain matrix equality problems in term of the clear and analytical expressions of the OLSEs and BLUEs obtained in the preceding sections. In order to resolve these proposed equivalence problems, we need to prepare a complex of preliminary methods and techniques in matrix algebra, including many formulas and facts about generalized inverses of matrices and the matrix rank methodology. Nevertheless, this sort of future investigations together with the current theoretical and methodological advances in this research area will definitely provide considerable insight into the intrinsic natures hidden behind the adding-up restrictions, so that we believe that the algebraic treatments presented in this article will sufficiently prompt other similar studies regarding different kinds of regression models with adding-up restrictions to observable random variables under various assumptions.

Acknowledgments

The author would like to express his sincere thanks to the handling editor and anonymous reviewers for their helpful comments and suggestions.

Conflict of interest

The author declares that there is no competing interest.

References

1. I. S. Alalouf, G. P. H. Styan, Characterizations of estimability in the general linear model, *Ann. Statist.*, **7** (1979), 194–200.
2. H. H. Bingham, W. J. Krzanowski, Linear algebra and multivariate analysis in statistics: development and interconnections in the twentieth century, *British J. Hist. Math.*, **37** (2022), 43–63. https://doi.org/10.1080/26375451.2022.2045811
3. H. Haupt, W. Oberhofer, Fully restricted linear regression: A pedagogical note, *Econ. Bull.*, **3** (2002), 1–7.

4. H. Haupt, W. Oberhofer, Generalized adding-up in systems of regression equations, *Econ. Lett.*, **92** (2006), 263–269. https://doi.org/10.1016/j.econlet.2006.03.001
5. A. Markiewicz, S. Puntanen, All about the $\perp$ with its applications in the linear statistical models, *Open Math.*, **13** (2015), 33–50. https://doi.org/10.1515/math-2015-0005
6. A. Markiewicz, S. Puntanen, G. P. H. Styan, The legend of the equality of OLSE and BLUE: highlighted by C. R. Rao in 1967, In: A volume in Honor of C. R. Rao on the occasion of his 100th birthday, *Methodol. Appl. Statist.*, 2021, 51–76.
7. G. Marsaglia, G. P. H. Styan, Equalities and inequalities for ranks of matrices, *Linear Multilinear Algebra*, **2** (1974), 269–292.
8. R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.*, **51** (1955), 406–413. https://doi.org/10.1017/S0305004100030401
9. S. Puntanen, G. P. H. Styan, J. Isotalo, *Matrix tricks for linear statistical models: our personal top twenty*, Berlin: Springer, 2011.
10. B. Ravikumar, S. Ray, N. E. Savin, Robust Wald tests in SUR systems with adding-up restrictions, *Econometrica*, **68** (2000), 715–719.
11. C. R. Rao, Unified theory of linear estimation, *Sankhyā Indian J. Statist. A*, **33** (1971), 371–394.
12. C. R. Rao, Representations of best linear unbiased estimators in the Gauss-Markoff model with a singular dispersion matrix, *J. Multivariate Anal.*, **3** (1973), 276–292. https://doi.org/10.1016/0047-259X(73)90042-0
13. M. Satchi, A note on adding-up restrictions when modelling trade flows, *Econ. Model.*, **21** (2004), 999–1002. https://doi.org/10.1016/j.econmod.2003.12.002
14. S. R. Searle, *Matrix algebra useful for statistics*, New York: Wiley, 1982.
15. Y. Tian, Some decompositions of OLSEs and BLUEs under a partitioned linear model, *Int. Stat. Rev.*, **75** (2007), 224–248. https://doi.org/10.1111/j.1751-5823.2007.00018.x
16. Y. Tian, On equalities of estimations of parametric functions under a general linear model and its restricted models, *Metrika*, **72** (2010), 313–330.
17. Y. Tian, Solving optimization problems on ranks and inertias of some constrained nonlinear matrix functions via an algebraic linearization method, *Nonlinear Anal.*, **75** (2012), 717–734. https://doi.org/10.1016/j.na.2011.09.003
18. Y. Tian, On properties of BLUEs under general linear regression models, *J. Statist. Plann. Inference*, **143** (2013), 771–782. https://doi.org/10.1016/j.jspi.2012.10.005
19. Y. Tian, M. Beisiegel, E. Dagenais, C. Haines, On the natural restrictions in the singular Gauss-Markov model, *Stat. Papers*, **49** (2007), 553–564. https://doi.org/10.1007/s00362-006-0032-5
20. Y. Tian, W. Guo, On comparison of dispersion matrices of estimators under a constrained linear model, *Stat. Methods Appl.*, **25** (2016), 623–649. https://doi.org/10.1007/s10260-016-0350-2
21. Y. Tian, J. Zhang, Some equalities for estimations of partial coefficients under a general linear regression model, *Stat. Papers*, **52** (2011), 911–920. https://doi.org/10.1007/s00362-009-0298-5