



Research article

A matrix analysis of BLMBPs under a general linear model and its transformation

Li Gong¹ and Bo Jiang^{2,*}

¹ Department of Primary School Education, Fuxin Higher Vocational College, Fuxin, Liaoning, China

² College of Mathematics and Information Science, Shandong Technology and Business University, Yantai, Shandong, China

* **Correspondence:** Email: jiangboliumengyu@gmail.com.

Abstract: This paper is concerned with the relationships between best linear minimum biased predictors (BLMBPs) in the context of a general linear model (GLM) and its transformed general linear models (TGLMs). We shall establish a mathematical procedure by means of some exact and analytical tools in matrix theory that were developed in recent years. The coverage includes constructing a general vector composed of all unknown parameters in the context of a GLM and its TGLMs, deriving the exact expressions of the BLMBPs through the technical use of analytical solutions of a constrained quadratic matrix-valued function optimization problem in the Löwner partial ordering, and discussing a variety of theoretical performances and properties of the BLMBPs. We also give a series of characterizations of relationships between BLMBPs under a given GLM and its TGLMs.

Keywords: general linear model; transformed general linear model; LMBP; BLMBP; relationships; rank

Mathematics Subject Classification: 15A09, 62H12, 62J05

1. Introduction

In this paper, we consider the standard linear model

$$\mathcal{M} : \mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\nu}, \tag{1.1}$$

where it is assumed that $\mathbf{y} \in \mathbb{R}^{n \times 1}$ is vector of observable random variables, $\mathbf{X} \in \mathbb{R}^{n \times p}$ are known matrices of arbitrary ranks ($0 \leq r(\mathbf{X}) \leq \min\{n, p\}$), $\boldsymbol{\theta} \in \mathbb{R}^{p \times 1}$ is a vector of fixed but unknown parameters, and $\boldsymbol{\nu} \in \mathbb{R}^{n \times 1}$ is a random error vector. In order to carry out reasonable estimation and

statistical inference in the context of (1.1), we assume that the expectation vector and the covariance matrix of \mathbf{v} are given without loss of generality by

$$E(\mathbf{v}) = \mathbf{0}, \quad \text{Cov}(\mathbf{v}) = \mathbf{\Sigma}. \quad (1.2)$$

In the sequel, we assume that $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ is a known positive semi-definite matrix of arbitrary rank in order to derive general and precise conclusions under the given model assumptions. Once the work on the general assumptions is established, we can further, as usual in parametric regression analysis, let $\mathbf{\Sigma}$ be certain specified forms with known or unknown entries, and then derive various concrete inference results.

The assumptions in the contexts of (1.1) and (1.2) are typical in form for a complete specification of the general linear model (for short, GLM). Observe that there are two unknown vectors $\boldsymbol{\theta}$ and \mathbf{v} in (1.1). Then, we can figure out that an obligatory task in statistical inference under (1.1) and (1.2) is to predict the two unknown vectors simultaneously:

$$\boldsymbol{\tau} = \mathbf{A}\boldsymbol{\theta} + \mathbf{B}\mathbf{v}, \quad (1.3)$$

where it is assumed that $\mathbf{A} \in \mathbb{R}^{s \times p}$ and $\mathbf{B} \in \mathbb{R}^{s \times n}$ are known matrices of arbitrary ranks. This vector obviously includes all the unknown vectors in (1.1), such as $\boldsymbol{\theta}$ and \mathbf{v} , as its special cases. It is easy to see that under (1.1) and (1.2), we have

$$E(\boldsymbol{\tau}) = \mathbf{A}\boldsymbol{\theta}, \quad \text{Cov}(\boldsymbol{\tau}) = \mathbf{B}\mathbf{\Sigma}\mathbf{B}', \quad \text{Cov}(\boldsymbol{\tau}, \mathbf{y}) = \mathbf{B}\mathbf{\Sigma}. \quad (1.4)$$

In the investigation of linear statistical models for regression, it is a common inference problem to propose and characterize various reasonable connections between two different given models under the given model assumptions. One concrete problem of such kind is to investigate the relationships between a given linear model (called the original model) and certain types of its transformed models. Sometimes, the transformed models are required to meet with certain necessary requirements in the statistical inferences of the original linear model. Now let us consider \mathcal{M} in (1.1) and its transformed models. In such a case, we may be faced with different transformed forms of the model in accordance with linear transformations of observable random vector \mathbf{y} . Generally speaking, various possible transformed models of \mathcal{M} in (1.1) are often obtained by pre-multiplying \mathbf{T} by a given matrix. For example,

$$\mathcal{N} : \quad \mathbf{T}\mathbf{y} = \mathbf{T}\mathbf{X}\boldsymbol{\theta} + \mathbf{T}\mathbf{v}, \quad (1.5)$$

is a common transformed form of \mathcal{M} in (1.1), where $\mathbf{T} \in \mathbb{R}^{m \times n}$ is a known transformation matrix of arbitrary rank. Below, we present a group of well-known cases of the transformed model for different choices of the transformation matrix \mathbf{T} in (1.5).

(a) We first divide the original model \mathcal{M} in (1.1) as

$$\mathcal{M} : \quad \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \boldsymbol{\theta} + \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}$$

by the partitions of the vectors and matrices in the model. Then we take the transformation matrices $\mathbf{T}_1 = [\mathbf{I}_{n_1}, \mathbf{0}]$ and $\mathbf{T}_2 = [\mathbf{0}, \mathbf{I}_{n_2}]$ in (1.5) to obtain the following two sub-sample models:

$$\mathcal{M}_1 : \quad \mathbf{y}_1 = \mathbf{X}_1\boldsymbol{\theta} + \mathbf{v}_1,$$

$$\mathcal{M}_2 : \mathbf{y}_2 = \mathbf{X}_2\boldsymbol{\theta} + \mathbf{v}_2,$$

where $\mathbf{y}_i \in \mathbb{R}^{n_i \times 1}$, $\mathbf{X}_i \in \mathbb{R}^{n_i \times p}$, $\boldsymbol{\theta} \in \mathbb{R}^{p \times 1}$, $\mathbf{v}_i \in \mathbb{R}^{n_i \times 1}$ and $n = n_1 + n_2$. They can also be viewed as adding or deleting certain regression equations in a given GLM. Also, we can say that these two individual models occur in two periods of observations of data.

(b) Assume that a concrete form of \mathcal{M} in (1.1) is given by

$$\mathcal{M} : \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}.$$

In this case, taking the two transformation matrices $\mathbf{T}_1 = [\mathbf{I}_{n_1}, \mathbf{0}]$ and $\mathbf{T}_2 = [\mathbf{0}, \mathbf{I}_{n_2}]$, we obtain the following two sub-sample models:

$$\mathcal{M}_1 : \mathbf{y}_1 = \mathbf{X}_1\boldsymbol{\theta}_1 + \mathbf{v}_1,$$

$$\mathcal{M}_2 : \mathbf{y}_2 = \mathbf{X}_2\boldsymbol{\theta}_2 + \mathbf{v}_2,$$

where $\mathbf{y}_i \in \mathbb{R}^{n_i \times 1}$, $\mathbf{X}_i \in \mathbb{R}^{n_i \times p_i}$, $\boldsymbol{\theta}_i \in \mathbb{R}^{p_i \times 1}$, $\mathbf{v}_i \in \mathbb{R}^{n_i \times 1}$, $n = n_1 + n_2$, $p = p_1 + p_2$. The two models are known as seemingly unrelated linear models, which are linked to each other by the correlated error terms across the models, where all the given matrices and the unknown vectors in the two models are different.

Due to the linear nature of \mathcal{M} in (1.1), we see the expectations and covariance matrices of \mathbf{y} , $\mathbf{T}\mathbf{y}$ and $\boldsymbol{\tau}$ under the assumptions in (1.1) and (1.2):

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\theta}, \quad E(\mathbf{T}\mathbf{y}) = \mathbf{T}\mathbf{X}\boldsymbol{\theta}, \quad (1.6)$$

$$\text{Cov}(\mathbf{y}) = \boldsymbol{\Sigma}, \quad \text{Cov}(\mathbf{T}\mathbf{y}) = \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}', \quad (1.7)$$

$$\text{Cov}(\boldsymbol{\tau}, \mathbf{y}) = \mathbf{B}\boldsymbol{\Sigma}, \quad \text{Cov}(\boldsymbol{\tau}, \mathbf{T}\mathbf{y}) = \mathbf{B}\boldsymbol{\Sigma}\mathbf{T}'. \quad (1.8)$$

Now we mention some backgrounds of this current study. For unknown parameters in a given regression model, statisticians are able to adopt different optimal criteria in order to obtain proper predictions and estimations of the unknown parameters. In comparison, the best linear unbiased prediction, the best linear unbiased estimation and the least squares estimation are best known among others because they have many excellent mathematical and statistical properties and performances. There were many deep and fruitful works in the statistical literatures related to these predictions and estimations. However, it is a common fact in statistical practice that the unknown parameters in a given model may not be predictable or estimable. Instead, it is necessary to choose certain biased predictions and biased estimations for the unknown parameters. For example, Rao described the bias between estimators and unknown parameter functions, constructed the minimum biased estimation class, selected the one with the minimum variance in the minimum biased estimation class and then defined the best linear minimum biased estimation. Especially when the unknown parameter function is an estimable function, the best linear minimum biased estimation is the classic best linear unbiased estimation. It can be seen from (1.1)–(1.8) that a given model and its transformed models are not necessarily equivalent in form. Hence, the predictors/estimators of unknown vectors that are going to derive under these models have different algebraic expressions and properties. Yet, some transformations of observable random vectors may preserve enough information for

predicting/estimating unknown vectors in the original model. Therefore, it is natural to consider certain links between the predictors/estimators obtained from an original model and its transformed models in statistical inferences of these models. Traditionally, the problems of characterizing relationships between predictions/estimations of unknown vectors in an original model and its transformed models were known as linear sufficiency problems, which were first considered in [1, 3]. Many scholars also studied the relationship between estimations under a given original model and its transformed model from different aspects. For instance, Baksalary and Kala considered the problem on linear transformations of GLMs preserving the best linear unbiased estimations under the general Gauss-Markoff model in [1]; Xie studied in [30] the best linear minimum biased estimations under a given GLM and discussed the problem of the linear transformation preserving the best linear minimum biased estimations. Also, the subject of this kind was sufficiently approached in [7, 9, 14, 18, 20, 33] among others.

Given the model assumptions in (1.1)–(1.8), the purpose of this paper is to provide a unified theoretical and conceptual exploration for solving the best linear minimum biased prediction (for short, BLMBP) problems under a GLM and its transformed general linear models (for short, TGLMs) through the skillful and effective use of a series of exact and analytical matrix analysis tools. The remaining part of this current paper is organized as follows. In the second section, we introduce notation and several matrix analysis tools and techniques that we shall utilize to characterize matrix equalities and matrix set inclusions that involve generalized inverses of matrices. In the third section, we introduce the definitions of the linear minimum biased predictor (for short, LMBP) and the BLMBP of τ in (1.3), as well as basic estimation and inference theory regarding the LMBP and BLMBP, including their analytical expressions and their mathematical and statistical properties and features in the contexts of (1.1)–(1.8). In the fourth section, we address the problems regarding the relationships between the BLMBPs under a GLM and its TGLMs using the powerful matrix rank and inertia methodology. The fifth section presents a special example related to the main findings in the preceding sections. Some conclusions and remarks are given in the last section.

2. Some preliminaries

We begin with the introduction of notation used in the sequel. $\mathbb{R}^{m \times n}$ denotes the collection of all $m \times n$ matrices over the field of real numbers, and the symbols \mathbf{M}' , $r(\mathbf{M})$ and $\mathcal{R}(\mathbf{M})$ denote the transpose, the rank and the range (column space) of $\mathbf{M} \in \mathbb{R}^{m \times n}$, and \mathbf{I}_m denotes the identity matrix of order m . The Moore–Penrose generalized inverse of \mathbf{M} , denoted by \mathbf{M}^\dagger , is defined to be the unique solution \mathbf{G} satisfying the four matrix equations $\mathbf{MGM} = \mathbf{M}$, $\mathbf{GMG} = \mathbf{G}$, $(\mathbf{MG})' = \mathbf{MG}$ and $(\mathbf{GM})' = \mathbf{GM}$. Let $\mathbf{P}_M = \mathbf{MM}^\dagger$, $\mathbf{M}^\perp = \mathbf{E}_M = \mathbf{I}_m - \mathbf{MM}^\dagger$ and $\mathbf{F}_M = \mathbf{I}_n - \mathbf{M}^\dagger\mathbf{M}$ denote the three orthogonal projectors (symmetric idempotent matrices) induced from \mathbf{M} , which will help in briefly denoting calculation processes related to generalized inverses of matrices, where both \mathbf{E}_M and \mathbf{F}_M satisfy $\mathbf{E}_M = \mathbf{F}_M'$ and $\mathbf{F}_M = \mathbf{E}_M'$ and the ranks of \mathbf{E}_M and \mathbf{F}_M are $r(\mathbf{E}_M) = m - r(\mathbf{M})$ and $r(\mathbf{F}_M) = n - r(\mathbf{M})$. Two symmetric matrices \mathbf{M} and \mathbf{N} of the same size are said to satisfy the inequalities $\mathbf{M} \succcurlyeq \mathbf{N}$, $\mathbf{M} \preccurlyeq \mathbf{N}$, $\mathbf{M} \succ \mathbf{N}$ and $\mathbf{M} \prec \mathbf{N}$ in the Löwner partial ordering if $\mathbf{M} - \mathbf{N}$ is positive semi-definite, negative semi-definite, positive definite and negative definite, respectively. Further information about the orthogonal projectors \mathbf{P}_M , \mathbf{E}_M and \mathbf{F}_M and their various applications in the theory of linear statistical models can be found, e.g., in [10, 13, 16, 19]. It is also well known that the Löwner partial ordering between two symmetric

matrices is a surprisingly strong and useful property in matrix analysis. The reader is referred to [16] and the references therein for more results and facts regarding the issues of the Löwner partial ordering in statistical theory and applications. Recently, the authors of [2, 4–6, 23, 25, 26, 29] proposed and approached a series of research problems concerning the relationships of different kinds of predictions of unknown parameters in regression models using the rank and inertia methodology in matrix analysis, and provided a variety of simple and reasonable equivalent facts related to the relationship problems. In this paper, we also adopt the rank and inertia methodology to approach the relationship problems regarding different estimations and predictions.

As preliminaries that can help readers in getting familiar with the features and usefulness of the matrix rank methodology, we present in the following a list of commonly used results and facts about ranks of matrices and matrix equations, which are well known or easy to prove. We shall use them in the descriptions and simplifications of various complicated matrix expressions and matrix equalities that occur in the statistical inference of a GLM and its TGLMs in the following sections.

Lemma 2.1 ([28]). *Let \mathcal{A} and \mathcal{B} be two sets composed by matrices of the same size.*

$$\mathcal{A} \cap \mathcal{B} \neq \emptyset \Rightarrow \min_{\mathbf{A} \in \mathcal{A}, \mathbf{B} \in \mathcal{B}} r(\mathbf{A} - \mathbf{B}) = 0, \quad (2.1)$$

$$\mathcal{A} \subseteq \mathcal{B} \Rightarrow \max_{\mathbf{A} \in \mathcal{A}} \min_{\mathbf{B} \in \mathcal{B}} r(\mathbf{A} - \mathbf{B}) = 0. \quad (2.2)$$

Lemma 2.2 ([12]). *Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$, and $\mathbf{C} \in \mathbb{R}^{l \times n}$. Then,*

$$r[\mathbf{A}, \mathbf{B}] = r(\mathbf{A}) + r(\mathbf{E}_A \mathbf{B}) = r(\mathbf{B}) + r(\mathbf{E}_B \mathbf{A}), \quad (2.3)$$

$$r \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) + r(\mathbf{C} \mathbf{F}_A) = r(\mathbf{C}) + r(\mathbf{A} \mathbf{F}_C), \quad (2.4)$$

$$r \begin{bmatrix} \mathbf{A} \mathbf{A}' & \mathbf{B} \\ \mathbf{B}' & \mathbf{0} \end{bmatrix} = r[\mathbf{A}, \mathbf{B}] + r(\mathbf{B}). \quad (2.5)$$

In particular, the following results hold:

$$(a) \quad r[\mathbf{A}, \mathbf{B}] = r(\mathbf{A}) \Leftrightarrow \mathcal{R}(\mathbf{B}) \subseteq \mathcal{R}(\mathbf{A}) \Leftrightarrow \mathbf{A} \mathbf{A}^\dagger \mathbf{B} = \mathbf{B} \Leftrightarrow \mathbf{E}_A \mathbf{B} = \mathbf{0}.$$

$$(b) \quad r \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} = r(\mathbf{A}) \Leftrightarrow \mathcal{R}(\mathbf{C}') \subseteq \mathcal{R}(\mathbf{A}') \Leftrightarrow \mathbf{C} \mathbf{A}^\dagger \mathbf{A} = \mathbf{C} \Leftrightarrow \mathbf{C} \mathbf{F}_A = \mathbf{0}.$$

Lemma 2.3 ([22]). *Assume that five matrices \mathbf{A}_1 , \mathbf{B}_1 , \mathbf{A}_2 , \mathbf{B}_2 and \mathbf{A}_3 of appropriate sizes satisfy the conditions $\mathcal{R}(\mathbf{A}'_1) \subseteq \mathcal{R}(\mathbf{B}'_1)$, $\mathcal{R}(\mathbf{A}_2) \subseteq \mathcal{R}(\mathbf{B}_1)$, $\mathcal{R}(\mathbf{A}'_2) \subseteq \mathcal{R}(\mathbf{B}'_2)$ and $\mathcal{R}(\mathbf{A}_3) \subseteq \mathcal{R}(\mathbf{B}_2)$. Then,*

$$r(\mathbf{A}_1 \mathbf{B}_1^\dagger \mathbf{A}_2 \mathbf{B}_2^\dagger \mathbf{A}_3) = r \begin{bmatrix} \mathbf{0} & \mathbf{B}_2 & \mathbf{A}_3 \\ \mathbf{B}_1 & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \end{bmatrix} - r(\mathbf{B}_1) - r(\mathbf{B}_2). \quad (2.6)$$

Lemma 2.4 ([21, 27]). *Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$ and $\mathbf{C} \in \mathbb{R}^{l \times n}$ be given. Then, the maximum and minimum ranks of $\mathbf{A} - \mathbf{B} \mathbf{Z}$ and $\mathbf{A} - \mathbf{B} \mathbf{Z} \mathbf{C}$ with respect to a variable matrix \mathbf{Z} of appropriate sizes are given by the following closed-form formulas:*

$$\max_{\mathbf{Z} \in \mathbb{R}^{k \times n}} r(\mathbf{A} - \mathbf{B} \mathbf{Z}) = \min\{r[\mathbf{A}, \mathbf{B}], n\}, \quad (2.7)$$

$$\min_{\mathbf{Z} \in \mathbb{R}^{k \times n}} r(\mathbf{A} - \mathbf{BZ}) = r[\mathbf{A}, \mathbf{B}] - r(\mathbf{B}), \quad (2.8)$$

$$\max_{\mathbf{Z} \in \mathbb{R}^{k \times l}} r(\mathbf{A} - \mathbf{BZC}) = \min \left\{ r[\mathbf{A}, \mathbf{B}], r \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \right\}. \quad (2.9)$$

Below we offer some existing formulas and results regarding general solutions of a basic linear matrix equation and a constrained quadratic matrix optimization problem.

Lemma 2.5 ([15]). *Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{B} \in \mathbb{R}^{p \times n}$. Then, the linear matrix equation $\mathbf{ZA} = \mathbf{B}$ is solvable for $\mathbf{Z} \in \mathbb{R}^{p \times m}$ if and only if $\mathcal{R}(\mathbf{A}') \supseteq \mathcal{R}(\mathbf{B}')$, or equivalently, $\mathbf{BA}^\dagger \mathbf{A} = \mathbf{B}$. In this case, the general solution of the equation can be written in the parametric form*

$$\mathbf{Z} = \mathbf{BA}^\dagger + \mathbf{UA}^\perp,$$

where $\mathbf{U} \in \mathbb{R}^{p \times m}$ is an arbitrary matrix.

Lemma 2.6 ([28]). *Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{R}^{m \times k}$ and assume that $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{B})$. Then*

$$\mathbf{XA} = \mathbf{0} \Leftrightarrow \mathbf{XB} = \mathbf{0}.$$

Lemma 2.7 ([24]). *Let*

$$f(\mathbf{Z}) = (\mathbf{ZC} + \mathbf{D})\mathbf{M}(\mathbf{ZC} + \mathbf{D})' \quad \text{s.t.} \quad \mathbf{ZA} = \mathbf{B}, \quad (2.10)$$

where it is assumed that $\mathbf{A} \in \mathbb{R}^{p \times q}$, $\mathbf{B} \in \mathbb{R}^{n \times q}$, $\mathbf{C} \in \mathbb{R}^{p \times m}$ and $\mathbf{D} \in \mathbb{R}^{n \times m}$ are given, $\mathbf{M} \in \mathbb{R}^{m \times m}$ is positive semi-definite and the matrix equation $\mathbf{ZA} = \mathbf{B}$ is solvable for $\mathbf{Z} \in \mathbb{R}^{n \times p}$. Then, there always exists a solution \mathbf{Z}_0 of $\mathbf{ZA} = \mathbf{B}$ such that

$$f(\mathbf{Z}) \geq f(\mathbf{Z}_0)$$

holds for all solutions of $\mathbf{ZA} = \mathbf{B}$, and the matrix \mathbf{Z}_0 that satisfies the above inequality is determined by the following consistent matrix equation:

$$\mathbf{Z}_0[\mathbf{A}, \mathbf{CMC}'\mathbf{A}^\perp] = [\mathbf{B}, -\mathbf{DMC}'\mathbf{A}^\perp].$$

In this case, the general expression of \mathbf{Z}_0 and the corresponding $f(\mathbf{Z}_0)$ and $f(\mathbf{Z})$ are given by

$$\begin{aligned} \mathbf{Z}_0 &= \underset{\mathbf{ZA}=\mathbf{B}}{\operatorname{argmin}} f(\mathbf{Z}) = [\mathbf{B}, -\mathbf{DMC}'\mathbf{A}^\perp][\mathbf{A}, \mathbf{CMC}'\mathbf{A}^\perp]^\dagger + \mathbf{U}[\mathbf{A}, \mathbf{CMC}']^\perp, \\ f(\mathbf{Z}_0) &= \underset{\mathbf{ZA}=\mathbf{B}}{\min} f(\mathbf{Z}) = \mathbf{GMG}' - \mathbf{GMC}'\mathbf{T}\mathbf{CMG}', \\ f(\mathbf{Z}) &= f(\mathbf{Z}_0) + (\mathbf{ZC} + \mathbf{D})\mathbf{MC}'\mathbf{T}\mathbf{CM}(\mathbf{ZC} + \mathbf{D})' \\ &= f(\mathbf{Z}_0) + (\mathbf{ZCMC}'\mathbf{A}^\perp + \mathbf{DMC}'\mathbf{A}^\perp)\mathbf{T}(\mathbf{ZCMC}'\mathbf{A}^\perp + \mathbf{DMC}'\mathbf{A}^\perp)', \end{aligned}$$

where $\mathbf{G} = \mathbf{BA}^\dagger \mathbf{C} + \mathbf{D}$, $\mathbf{T} = (\mathbf{A}^\perp \mathbf{CMC}' \mathbf{A}^\perp)^\dagger$ and $\mathbf{U} \in \mathbb{R}^{n \times p}$ is arbitrary.

In order to describe the relationships between BLMBPs under different regression models, we need to adopt the following definition to characterize possible equality between two random vectors [28].

Definition 2.8. *Let \mathbf{y} be as given in (1.1), let $\{\mathbf{L}_1\}$ and $\{\mathbf{L}_2\}$ be two matrix sets and let $\mathbf{L}_1\mathbf{y}$ and $\mathbf{L}_2\mathbf{y}$ be any two linear predictors of $\boldsymbol{\tau}$ in (1.3).*

(a) $\{\mathbf{L}_1\mathbf{y}\} \cap \{\mathbf{L}_2\mathbf{y}\} \neq \emptyset$ holds definitely, i.e. $\{\mathbf{L}_1\} \cap \{\mathbf{L}_2\} \neq \emptyset$ if and only if

$$\min_{\mathbf{L}_1 \in \{\mathbf{L}_1\}, \mathbf{L}_2 \in \{\mathbf{L}_2\}} r(\mathbf{L}_1 - \mathbf{L}_2) = 0.$$

(b) The vector set inclusion $\{\mathbf{L}_1\mathbf{y}\} \subseteq \{\mathbf{L}_2\mathbf{y}\}$ holds definitely, i.e. $\{\mathbf{L}_1\} \subseteq \{\mathbf{L}_2\}$ if and only if

$$\max_{\mathbf{L}_1 \in \{\mathbf{L}_1\}} \min_{\mathbf{L}_2 \in \{\mathbf{L}_2\}} r(\mathbf{L}_1 - \mathbf{L}_2) = 0.$$

(c) $\{\mathbf{L}_1\mathbf{y}\} \cap \{\mathbf{L}_2\mathbf{y}\} \neq \emptyset$ holds with probability 1 if and only if

$$\begin{aligned} & \min_{\mathbf{L}_1 \in \{\mathbf{L}_1\}, \mathbf{L}_2 \in \{\mathbf{L}_2\}} r((\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}, \boldsymbol{\Sigma}]) = 0 \\ \Leftrightarrow & \min_{\mathbf{L}_1 \in \{\mathbf{L}_1\}, \mathbf{L}_2 \in \{\mathbf{L}_2\}} r((\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}, \boldsymbol{\Sigma}\mathbf{X}^\perp]) = 0 \\ \Leftrightarrow & \min_{\mathbf{L}_1 \in \{\mathbf{L}_1\}, \mathbf{L}_2 \in \{\mathbf{L}_2\}} r((\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) = 0. \end{aligned}$$

(d) The vector set inclusion $\{\mathbf{L}_1\mathbf{y}\} \subseteq \{\mathbf{L}_2\mathbf{y}\}$ holds with probability 1 if and only if

$$\begin{aligned} & \max_{\mathbf{L}_1 \in \{\mathbf{L}_1\}} \min_{\mathbf{L}_2 \in \{\mathbf{L}_2\}} r((\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}, \boldsymbol{\Sigma}]) = 0 \\ \Leftrightarrow & \max_{\mathbf{L}_1 \in \{\mathbf{L}_1\}} \min_{\mathbf{L}_2 \in \{\mathbf{L}_2\}} r((\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}, \boldsymbol{\Sigma}\mathbf{X}^\perp]) = 0 \\ \Leftrightarrow & \max_{\mathbf{L}_1 \in \{\mathbf{L}_1\}} \min_{\mathbf{L}_2 \in \{\mathbf{L}_2\}} r((\mathbf{L}_1 - \mathbf{L}_2)[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) = 0. \end{aligned}$$

3. Fundamentals of the LMBP/LMBE and BLMBP/BLMBE

Recall in parametric regression analysis that if there exists a matrix \mathbf{L} such that $\mathbf{E}(\mathbf{L}\mathbf{y} - \boldsymbol{\tau}) = (\mathbf{L}\mathbf{X} - \mathbf{A})\boldsymbol{\theta} = \mathbf{0}$ holds for all $\boldsymbol{\theta}$, the parametric parameter vector $\boldsymbol{\tau}$ in (1.3) is said to be predictable under the assumptions in (1.1) and (1.2). Otherwise, there does not exist an unbiased prediction of $\boldsymbol{\tau}$ under (1.1) and (1.2), and therefore, we have to seek certain biased predictions of $\boldsymbol{\tau}$ according to various specified optimization criteria. In this section, we shall adopt the following known definitions of the LMBP and BLMBP of $\boldsymbol{\tau}$ (cf. [17, p. 337]).

Definition 3.1. Let the parametric vector $\boldsymbol{\tau}$ be as given in (1.3).

(a) The LMBP of $\boldsymbol{\tau}$ in (1.3) under (1.1) is defined to be

$$\text{LMBP}_{\mathcal{M}}(\boldsymbol{\tau}) = \text{LMBP}_{\mathcal{M}}(\mathbf{A}\boldsymbol{\theta} + \mathbf{B}\boldsymbol{\nu}) = \widehat{\mathbf{L}}\mathbf{y}, \quad (3.1)$$

where the matrix $\widehat{\mathbf{L}}$ satisfies

$$\widehat{\mathbf{L}} = \underset{\mathbf{L} \in \mathbb{R}^{s \times n}}{\text{argmin}} \text{tr}((\mathbf{L}\mathbf{X} - \mathbf{A})(\mathbf{L}\mathbf{X} - \mathbf{A})'). \quad (3.2)$$

(b) The LMBP of $\boldsymbol{\tau}$ in (1.3) under (1.5) is defined to be

$$\text{LMBP}_{\mathcal{N}}(\boldsymbol{\tau}) = \text{LMBP}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta} + \mathbf{B}\boldsymbol{\nu}) = \widehat{\mathbf{K}}\mathbf{T}\mathbf{y}, \quad (3.3)$$

where the matrix $\widehat{\mathbf{K}}$ satisfies

$$\widehat{\mathbf{K}} = \underset{\mathbf{K} \in \mathbb{R}^{s \times m}}{\text{argmin}} \text{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})'). \quad (3.4)$$

Theorem 3.2. Under the notations in Definition 3.1, the following results hold:

$$\widehat{\mathbf{L}} = \operatorname{argmin}_{\mathbf{L} \in \mathbb{R}^{s \times n}} \operatorname{tr}((\mathbf{L}\mathbf{X} - \mathbf{A})(\mathbf{L}\mathbf{X} - \mathbf{A})') \Leftrightarrow \widehat{\mathbf{L}}\mathbf{X}\mathbf{X}' = \mathbf{A}\mathbf{X}', \quad (3.5)$$

$$\widehat{\mathbf{K}} = \operatorname{argmin}_{\mathbf{K} \in \mathbb{R}^{s \times m}} \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})') \Leftrightarrow \widehat{\mathbf{K}}\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})' = \mathbf{A}(\mathbf{T}\mathbf{X})'. \quad (3.6)$$

Proof. Note first that

$$\begin{aligned} & \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})') \\ &= \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} + \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} + \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} - \mathbf{A})') \\ &= \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} - \mathbf{A}\mathbf{F}_{\mathbf{T}\mathbf{X}})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} - \mathbf{A}\mathbf{F}_{\mathbf{T}\mathbf{X}})') \\ &= \operatorname{tr}(\mathbf{A}\mathbf{F}_{\mathbf{T}\mathbf{X}}\mathbf{A}') + \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X})') \\ &\quad - \operatorname{tr}((\mathbf{K} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger)\mathbf{T}\mathbf{X}\mathbf{F}_{\mathbf{T}\mathbf{X}}\mathbf{A}') - \operatorname{tr}(\mathbf{A}\mathbf{F}_{\mathbf{T}\mathbf{X}}(\mathbf{T}\mathbf{X})'(\mathbf{K} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger)') \\ &= \operatorname{tr}(\mathbf{A}\mathbf{F}_{\mathbf{T}\mathbf{X}}\mathbf{A}') + \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X})'), \end{aligned} \quad (3.7)$$

where $\mathbf{T}\mathbf{X}\mathbf{F}_{\mathbf{T}\mathbf{X}} = \mathbf{0}$. Note that $\operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X})') \geq 0$ for all $\mathbf{K} \in \mathbb{R}^{s \times m}$ and the matrix equation $\mathbf{K}\mathbf{T}\mathbf{X} = \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X}$ is solvable for $\mathbf{K} \in \mathbb{R}^{s \times m}$. In this case, we obtain

$$\min_{\mathbf{K} \in \mathbb{R}^{s \times m}} \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})') = \operatorname{tr}(\mathbf{A}\mathbf{F}_{\mathbf{T}\mathbf{X}}\mathbf{A}'),$$

and

$$\widehat{\mathbf{K}} = \operatorname{argmin}_{\mathbf{K} \in \mathbb{R}^{s \times m}} \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})') \Leftrightarrow \widehat{\mathbf{K}}\mathbf{T}\mathbf{X} - \mathbf{A}(\mathbf{T}\mathbf{X})^\dagger\mathbf{T}\mathbf{X} = \mathbf{0} \Leftrightarrow \widehat{\mathbf{K}}\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})' = \mathbf{A}(\mathbf{T}\mathbf{X})',$$

thus establishing (3.6). Letting $\mathbf{T} = \mathbf{I}_n$ leads to (3.5). \square

Definition 3.3. Let the parametric vector $\boldsymbol{\tau}$ be as given in (1.3).

(a) If $\widehat{\mathbf{L}}$ satisfies

$$\operatorname{Cov}(\widehat{\mathbf{L}}\mathbf{y} - \boldsymbol{\tau}) = \min \quad \text{s.t.} \quad \widehat{\mathbf{L}} = \operatorname{argmin}_{\mathbf{L} \in \mathbb{R}^{s \times n}} \operatorname{tr}((\mathbf{L}\mathbf{X} - \mathbf{A})(\mathbf{L}\mathbf{X} - \mathbf{A})') \quad (3.8)$$

holds in the Löwner partial ordering, then the linear statistic $\widehat{\mathbf{L}}\mathbf{y}$ is defined to be the BLMBP of $\boldsymbol{\tau}$ in (1.3) under (1.1), and is denoted by

$$\widehat{\mathbf{L}}\mathbf{y} = \operatorname{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau}) = \operatorname{BLMBP}_{\mathcal{M}}(\mathbf{A}\boldsymbol{\theta} + \mathbf{B}\boldsymbol{\nu}). \quad (3.9)$$

(b) If $\widehat{\mathbf{K}}$ satisfies

$$\operatorname{Cov}(\widehat{\mathbf{K}}\mathbf{T}\mathbf{y} - \boldsymbol{\tau}) = \min \quad \text{s.t.} \quad \widehat{\mathbf{K}} = \operatorname{argmin}_{\mathbf{K} \in \mathbb{R}^{s \times m}} \operatorname{tr}((\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})(\mathbf{K}\mathbf{T}\mathbf{X} - \mathbf{A})') \quad (3.10)$$

holds in the Löwner partial ordering, then the linear statistic $\widehat{\mathbf{K}}\mathbf{T}\mathbf{y}$ is defined to be the BLMBP of $\boldsymbol{\tau}$ in (1.3) under (1.5), and is denoted by

$$\widehat{\mathbf{K}}\mathbf{T}\mathbf{y} = \operatorname{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}) = \operatorname{BLMBP}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta} + \mathbf{B}\boldsymbol{\nu}). \quad (3.11)$$

If $\mathbf{B} = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$ in (1.3), then the $\widehat{\mathbf{K}}\mathbf{T}\mathbf{y}$ in (3.11) are defined to be the best linear minimum biased estimator (BLMBE) and the BLMBP of $\mathbf{A}\boldsymbol{\theta}$ and $\mathbf{B}\boldsymbol{\nu}$ in (1.3) under (1.5), respectively, and are denoted by

$$\widehat{\mathbf{K}}\mathbf{T}\mathbf{y} = \operatorname{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}) \quad \text{and} \quad \widehat{\mathbf{K}}\mathbf{T}\mathbf{y} = \operatorname{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu}).$$

It is easy to verify that the difference of $\mathbf{KTy} - \boldsymbol{\tau}$ under (1.5) can be written in the following form:

$$\mathbf{KTy} - \boldsymbol{\tau} = \mathbf{KTX}\boldsymbol{\theta} + \mathbf{KT}\boldsymbol{\nu} - \mathbf{A}\boldsymbol{\theta} - \mathbf{B}\boldsymbol{\nu} = (\mathbf{KTX} - \mathbf{A})\boldsymbol{\theta} + (\mathbf{KT} - \mathbf{B})\boldsymbol{\nu}.$$

Hence, the covariance matrix of $\mathbf{KTy} - \boldsymbol{\tau}$ can be written as

$$\text{Cov}(\mathbf{KTy} - \boldsymbol{\tau}) = (\mathbf{KT} - \mathbf{B})\boldsymbol{\Sigma}(\mathbf{KT} - \mathbf{B})' \triangleq f(\mathbf{K}). \quad (3.12)$$

Our main results on the BLMBPs of $\boldsymbol{\tau}$ in (1.3) are given below.

Theorem 3.4. *Let the parametric vector $\boldsymbol{\tau}$ be as given in (1.3) and define $\mathbf{W} = [\mathbf{TX}(\mathbf{TX})', \text{Cov}(\mathbf{Ty})(\mathbf{TX})^\perp]$ and $\mathbf{D} = \text{Cov}(\boldsymbol{\tau}, \mathbf{Ty})$. Then*

$$\text{Cov}(\widehat{\mathbf{K}}\mathbf{Ty} - \boldsymbol{\tau}) = \min \text{ s.t. } \widehat{\mathbf{K}}\mathbf{TX}(\mathbf{TX})' = \mathbf{A}(\mathbf{TX})' \Leftrightarrow \widehat{\mathbf{K}}\mathbf{W} = [\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]. \quad (3.13)$$

The matrix equation in (3.13) is solvable for $\widehat{\mathbf{K}}$, i.e.,

$$[\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger\mathbf{W} = [\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp] \quad (3.14)$$

holds under (3.6), while the general expression of $\widehat{\mathbf{K}}$ and the corresponding $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ can be written in the following form

$$\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}) = \widehat{\mathbf{K}}\mathbf{Ty} = ([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger + \mathbf{U}_1\mathbf{W}^\perp)\mathbf{Ty}, \quad (3.15)$$

where $\mathbf{U}_1 \in \mathbb{R}^{s \times m}$ is arbitrary. Furthermore, the following results hold.

- (a) $r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] = r[\mathbf{TX}, (\mathbf{TX})^\perp\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'] = r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}]$ and $\mathcal{R}[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] = \mathcal{R}[\mathbf{TX}, (\mathbf{TX})^\perp\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'] = \mathcal{R}[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}]$.
- (b) $\widehat{\mathbf{K}}\mathbf{T}$ in (3.15) is unique if and only if $\mathcal{R}(\mathbf{T}) \subseteq \mathcal{R}[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}]$.
- (c) $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ is unique if and only if $\mathbf{Ty} \in \mathcal{R}[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}]$ holds with probability 1.
- (d) The covariance matrix of $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ is given by

$$\text{Cov}(\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})) = \widehat{\mathbf{K}}\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'\widehat{\mathbf{K}}' = ([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger)\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger)'; \quad (3.16)$$

the covariance matrix between $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ and $\boldsymbol{\tau}$ is given by

$$\text{Cov}(\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}), \boldsymbol{\tau}) = [\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp][\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\dagger\mathbf{D}'; \quad (3.17)$$

the difference of $\text{Cov}(\boldsymbol{\tau})$ and $\text{Cov}(\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}))$ is given by

$$\text{Cov}(\boldsymbol{\tau}) - \text{Cov}(\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})) = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}' - ([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger)\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger)'; \quad (3.18)$$

the covariance matrix of $\boldsymbol{\tau} - \text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ is given by

$$\text{Cov}(\boldsymbol{\tau} - \text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})) = ([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger\mathbf{T} - \mathbf{B})\boldsymbol{\Sigma}([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger\mathbf{T} - \mathbf{B})'. \quad (3.19)$$

(e) If $\mathbf{B} = \mathbf{0}$ or $\mathbf{A} = \mathbf{0}$ in (1.3), then

$$\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}) = ([\mathbf{A}(\mathbf{TX})', \mathbf{0}]\mathbf{W}^\dagger + \mathbf{U}_1\mathbf{W}^\perp)\mathbf{T}\mathbf{y}, \quad (3.20)$$

$$\text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu}) = ([\mathbf{0}, \mathbf{D}(\mathbf{TX})^\perp]\mathbf{W}^\dagger + \mathbf{U}_1\mathbf{W}^\perp)\mathbf{T}\mathbf{y}. \quad (3.21)$$

Proof. Eq (3.13) is obviously equivalent to

$$f(\mathbf{K}) = (\mathbf{KT} - \mathbf{B})\boldsymbol{\Sigma}(\mathbf{KT} - \mathbf{B})' = \min \quad \text{s.t.} \quad \mathbf{KTX}(\mathbf{TX})' = \mathbf{A}(\mathbf{TX})'. \quad (3.22)$$

Since $\boldsymbol{\Sigma} \succeq \mathbf{0}$, the optimization problem in (3.22) is a special case of (2.10). By Lemma 2.7, the solution of (3.22) is determined by the matrix equation in (3.13). This equation is consistent as well under (3.6), and the general solution of the equation and the corresponding BLMBP are given in (3.15). Result (a) is well known; see [11, 16]. Results (b) and (c) follow from the conditions $[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} = \mathbf{0}$ and $[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}']^\perp \mathbf{T}\mathbf{y} = \mathbf{0}$ holds with probability 1.

Taking the covariance operation of (3.15) yields (3.16). Also from (1.8) and (3.15), the covariance matrix between $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ and $\boldsymbol{\tau}$ is

$$\begin{aligned} \text{Cov}(\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}), \boldsymbol{\tau}) &= \text{Cov}(\widehat{\mathbf{K}}\mathbf{T}\mathbf{y}, \boldsymbol{\tau}) \\ &= [\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp][\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\dagger \mathbf{T}\boldsymbol{\Sigma}\mathbf{B}' \\ &= [\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp][\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\dagger \mathbf{D}', \end{aligned}$$

thus establishing (3.17). Combination of (1.4) and (3.16) yields (3.18). Substitution of (3.15) into (3.12) and then simplification yields (3.19). \square

Some conclusions for a special case of Theorem 3.4 are presented below without proof.

Corollary 3.5. *Let the parametric vector $\boldsymbol{\tau}$ be as given in (1.3), and define $\mathbf{V} = [\mathbf{X}\mathbf{X}', \text{Cov}(\mathbf{y})\mathbf{X}^\perp]$ and $\mathbf{C} = \text{Cov}(\boldsymbol{\tau}, \mathbf{y})$. Then*

$$\text{Cov}(\widehat{\mathbf{L}}\mathbf{y} - \boldsymbol{\tau}) = \min \quad \text{s.t.} \quad \widehat{\mathbf{L}}\mathbf{X}\mathbf{X}' = \mathbf{A}\mathbf{X}' \Leftrightarrow \widehat{\mathbf{L}}\mathbf{V} = [\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^\perp]. \quad (3.23)$$

The matrix equation in (3.23) is solvable for $\widehat{\mathbf{L}}$, i.e.,

$$[\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^\perp]\mathbf{V}^\dagger\mathbf{V} = [\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^\perp] \quad (3.24)$$

holds under (3.5), while the general expression of $\widehat{\mathbf{L}}$ and the corresponding $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ can be written in the following form

$$\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau}) = \widehat{\mathbf{L}}\mathbf{y} = ([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^\perp]\mathbf{V}^\dagger + \mathbf{U}_2\mathbf{V}^\perp)\mathbf{y}, \quad (3.25)$$

where $\mathbf{U}_2 \in \mathbb{R}^{s \times n}$ is arbitrary. Furthermore, the following results hold.

- (a) $r[\mathbf{X}, \boldsymbol{\Sigma}\mathbf{X}^\perp] = r[\mathbf{X}, \mathbf{X}^\perp\boldsymbol{\Sigma}] = r[\mathbf{X}, \boldsymbol{\Sigma}]$ and $\mathcal{R}[\mathbf{X}, \boldsymbol{\Sigma}\mathbf{X}^\perp] = \mathcal{R}[\mathbf{X}, \mathbf{X}^\perp\boldsymbol{\Sigma}] = \mathcal{R}[\mathbf{X}, \boldsymbol{\Sigma}]$.
- (b) $\widehat{\mathbf{L}}$ in (3.25) is unique if and only if $r[\mathbf{X}, \boldsymbol{\Sigma}] = n$.
- (c) $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ is unique if and only if $\mathbf{y} \in \mathcal{R}[\mathbf{X}, \boldsymbol{\Sigma}]$ holds with probability 1.

(d) The covariance matrix of $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ is given by

$$\text{Cov}(\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})) = \widehat{\mathbf{L}}\widehat{\boldsymbol{\Sigma}}\widehat{\mathbf{L}}' = ([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}]\mathbf{V}^{\dagger})\boldsymbol{\Sigma}([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}]\mathbf{V}^{\dagger})'; \quad (3.26)$$

the covariance matrix between $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ and $\boldsymbol{\tau}$ is given by

$$\text{Cov}(\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau}), \boldsymbol{\tau}) = [\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}][\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^{\perp}]^{\dagger}\mathbf{C}'; \quad (3.27)$$

the difference of $\text{Cov}(\boldsymbol{\tau})$ and $\text{Cov}(\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau}))$ is given by

$$\text{Cov}(\boldsymbol{\tau}) - \text{Cov}(\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})) = \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}' - ([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}]\mathbf{V}^{\dagger})\boldsymbol{\Sigma}([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}]\mathbf{V}^{\dagger})'; \quad (3.28)$$

the covariance matrix of $\boldsymbol{\tau} - \text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ is given by

$$\text{Cov}(\boldsymbol{\tau} - \text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})) = ([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}]\mathbf{V}^{\dagger} - \mathbf{B})\boldsymbol{\Sigma}([\mathbf{A}\mathbf{X}', \mathbf{C}\mathbf{X}^{\perp}]\mathbf{V}^{\dagger} - \mathbf{B})'. \quad (3.29)$$

Corollary 3.6. Let the parametric vector $\boldsymbol{\tau}$ be as given in (1.3). Then, the following results hold:

(a) The BLMBP of $\boldsymbol{\tau}$ can be decomposed as the sum

$$\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}) = \text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}) + \text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu}), \quad (3.30)$$

and they satisfy

$$\text{Cov}(\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}), \text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu})) = \mathbf{0}, \quad (3.31)$$

$$\text{Cov}(\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})) = \text{Cov}(\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta})) + \text{Cov}(\text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu})). \quad (3.32)$$

(b) For any matrix $\mathbf{P} \in \mathbb{R}^{t \times s}$, and the following equality holds:

$$\text{BLMBP}_{\mathcal{N}}(\mathbf{P}\boldsymbol{\tau}) = \mathbf{P}\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}). \quad (3.33)$$

(c) The BLMBP of $\boldsymbol{\tau}$ can be decomposed as the sum

$$\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau}) = \text{BLMBE}_{\mathcal{M}}(\mathbf{A}\boldsymbol{\theta}) + \text{BLMBP}_{\mathcal{M}}(\mathbf{B}\boldsymbol{\nu}), \quad (3.34)$$

and they satisfy

$$\text{Cov}(\text{BLMBE}_{\mathcal{M}}(\mathbf{A}\boldsymbol{\theta}), \text{BLMBP}_{\mathcal{M}}(\mathbf{B}\boldsymbol{\nu})) = \mathbf{0}, \quad (3.35)$$

$$\text{Cov}(\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})) = \text{Cov}(\text{BLMBE}_{\mathcal{M}}(\mathbf{A}\boldsymbol{\theta})) + \text{Cov}(\text{BLMBP}_{\mathcal{M}}(\mathbf{B}\boldsymbol{\nu})). \quad (3.36)$$

(d) For any matrix $\mathbf{P} \in \mathbb{R}^{t \times s}$, and the following equality holds:

$$\text{BLMBP}_{\mathcal{M}}(\mathbf{P}\boldsymbol{\tau}) = \mathbf{P}\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau}). \quad (3.37)$$

Proof. Notice that the arbitrary matrix \mathbf{U}_1 in (3.15) can be rewritten as $\mathbf{U}_1 = \mathbf{V}_1 + \mathbf{V}_2$, while the matrix $[\mathbf{A}(\mathbf{T}\mathbf{X})', \mathbf{D}(\mathbf{T}\mathbf{X})^{\perp}]$ in (3.15) can be rewritten as

$$[\mathbf{A}(\mathbf{T}\mathbf{X})', \mathbf{D}(\mathbf{T}\mathbf{X})^{\perp}] = [\mathbf{A}(\mathbf{T}\mathbf{X})', \mathbf{0}] + [\mathbf{0}, \mathbf{D}(\mathbf{T}\mathbf{X})^{\perp}].$$

Correspondingly, $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ in (3.15) can be rewritten as the sum:

$$\begin{aligned}\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau}) &= ([\mathbf{A}(\mathbf{TX})', \mathbf{D}(\mathbf{TX})^\perp] \mathbf{W}^\dagger + \mathbf{U}_1 \mathbf{W}^\perp) \mathbf{T} \mathbf{y} \\ &= ([\mathbf{A}(\mathbf{TX})', \mathbf{0}] \mathbf{W}^\dagger + \mathbf{V}_1 \mathbf{W}^\perp) \mathbf{T} \mathbf{y} \\ &\quad + ([\mathbf{0}, \mathbf{D}(\mathbf{TX})^\perp] \mathbf{W}^\dagger + \mathbf{V}_2 \mathbf{W}^\perp) \mathbf{T} \mathbf{y} \\ &= \text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}) + \text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu}),\end{aligned}$$

thus establishing (3.30). From (3.20) and (3.21), the covariance matrix between $\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta})$ and $\text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu})$ is given by

$$\text{Cov}(\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}), \text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu})) = [\mathbf{A}(\mathbf{TX})', \mathbf{0}] \mathbf{W}^\dagger \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' ([\mathbf{0}, \mathbf{B} \boldsymbol{\Sigma} \mathbf{T}' (\mathbf{TX})^\perp] \mathbf{W}^\dagger)'$$

Applying (2.6) to the right-hand side of the above equality and then simplifying by Theorem 3.4(a), (2.3), and (2.5), we obtain

$$\begin{aligned}& r(\text{Cov}(\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}), \text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu}))) \\ &= r([\mathbf{A}(\mathbf{TX})', \mathbf{0}] \mathbf{W}^\dagger \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' ([\mathbf{0}, \mathbf{B} \boldsymbol{\Sigma} \mathbf{T}' (\mathbf{TX})^\perp] \mathbf{W}^\dagger)') \\ &= r \left[\begin{array}{cc} \mathbf{0} & \begin{bmatrix} \mathbf{TX}(\mathbf{TX})' \\ (\mathbf{TX})^\perp \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' \end{bmatrix} \\ \begin{bmatrix} \mathbf{TX}(\mathbf{TX})' & \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' \\ \mathbf{A}(\mathbf{TX})' & \mathbf{0} \end{bmatrix} & \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \end{array} \right] - 2r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' (\mathbf{TX})^\perp] \\ &= r \left[\begin{array}{cc} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -(\mathbf{TX})^\perp \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' (\mathbf{TX})^\perp \end{bmatrix} & \begin{bmatrix} \mathbf{TX}(\mathbf{TX})' \\ \mathbf{0} \end{bmatrix} \\ \begin{bmatrix} \mathbf{TX}(\mathbf{TX})' & \mathbf{0} \\ \mathbf{A}(\mathbf{TX})' & \mathbf{0} \end{bmatrix} & \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \end{array} \right] - 2r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}] \\ &= r \left[\begin{array}{cc} \mathbf{0} & \mathbf{TX}(\mathbf{TX})' \\ \mathbf{TX}(\mathbf{TX})' & \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' \\ \mathbf{A}(\mathbf{TX})' & \mathbf{0} \end{array} \right] + r[(\mathbf{TX})^\perp \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' (\mathbf{TX})^\perp, (\mathbf{TX})^\perp \mathbf{T} \boldsymbol{\Sigma} \mathbf{B}'] - 2r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}] \\ &= r \left[\begin{array}{c} \mathbf{TX}(\mathbf{TX})' \\ \mathbf{A}(\mathbf{TX})' \end{array} \right] + r \left[\begin{array}{c} \mathbf{TX}(\mathbf{TX})' \\ \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' \end{array} \right] + r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma} \mathbf{T}' (\mathbf{TX})^\perp, \mathbf{T} \boldsymbol{\Sigma} \mathbf{B}'] - r(\mathbf{TX}) - 2r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}] \\ &= r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}, \mathbf{T} \boldsymbol{\Sigma} \mathbf{B}'] - r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}] \\ &= r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T} \boldsymbol{\Sigma}] \\ &= 0,\end{aligned}$$

which implies that $\text{Cov}(\text{BLMBE}_{\mathcal{N}}(\mathbf{A}\boldsymbol{\theta}), \text{BLMBP}_{\mathcal{N}}(\mathbf{B}\boldsymbol{\nu}))$ is a zero matrix, thus establishing (3.31). Equation (3.32) follows from (3.30) and (3.31). Result (b) follows directly from (3.15). Results (c) and (d) are special cases of (a) and (b). \square

4. Relationships between the BLMBPs under a GLM and its TGLMs

One of the main tasks in the statistical inference of parametric regression models is to characterize connections between different predictions/estimations of unknown parameters. In this section, we

study the relationships between the BLMBPs under GLM and its TGLMs. Because the coefficient matrices $\widehat{\mathbf{K}}\mathbf{T}$ and $\widehat{\mathbf{L}}$ in (3.15) and (3.25) are not necessarily unique, we use

$$\{\widehat{\mathbf{L}}\}, \{\widehat{\mathbf{K}}\mathbf{T}\}, \{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} = \{\widehat{\mathbf{L}}\mathbf{y}\}, \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} = \{\widehat{\mathbf{K}}\mathbf{T}\mathbf{y}\} \tag{4.1}$$

to denote the collections of all the coefficient matrices and the corresponding BLMBPs. In order to characterize the relations between the collections of the coefficient matrices in (4.1), it is necessary to discuss the following four cases:

- (a) $\{\widehat{\mathbf{L}}\} \cap \{\widehat{\mathbf{K}}\mathbf{T}\} \neq \emptyset$, so that $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds definitely;
- (b) $\{\widehat{\mathbf{L}}\} \supseteq \{\widehat{\mathbf{K}}\mathbf{T}\}$, so that $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \supseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely;
- (c) $\{\widehat{\mathbf{L}}\} \subseteq \{\widehat{\mathbf{K}}\mathbf{T}\}$, so that $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \subseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely;
- (d) $\{\widehat{\mathbf{L}}\} = \{\widehat{\mathbf{K}}\mathbf{T}\}$, so that $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} = \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely.

In order to characterize the relations between the collections of the random vectors in (4.1), it is necessary to discuss the following four cases:

- (a) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds with probability 1;
- (b) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \supseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1;
- (c) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \subseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1;
- (d) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} = \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.

Our main results are given below.

Theorem 4.1. Let $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ and $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ be as given in (3.15) and (3.25), respectively and define

$$\boldsymbol{\Lambda} = \begin{bmatrix} \mathbf{T}\mathbf{X}\mathbf{X}' & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \end{bmatrix}, \boldsymbol{\Gamma} = [\mathbf{A}\mathbf{X}', \mathbf{B}\boldsymbol{\Sigma}]. \tag{4.2}$$

Then, the following results hold.

- (a) There exist $\widehat{\mathbf{L}}$ and $\widehat{\mathbf{K}}$ such that $\widehat{\mathbf{L}} = \widehat{\mathbf{K}}\mathbf{T}$ if and only if $\mathcal{R}(\boldsymbol{\Gamma}') \subseteq \mathcal{R}(\boldsymbol{\Lambda}')$. In this case, $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds definitely.
- (b) $\{\widehat{\mathbf{L}}\} \supseteq \{\widehat{\mathbf{K}}\mathbf{T}\}$ if and only if $\mathcal{R}(\boldsymbol{\Gamma}') \subseteq \mathcal{R}(\boldsymbol{\Lambda}')$. In this case, $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \supseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely.
- (c) $\{\widehat{\mathbf{L}}\} \subseteq \{\widehat{\mathbf{K}}\mathbf{T}\}$ if and only if $r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} = r(\mathbf{T}) + r(\mathbf{X}) + r[\mathbf{X}, \boldsymbol{\Sigma}] - n$. In this case, $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \subseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely.
- (d) $\{\widehat{\mathbf{L}}\} = \{\widehat{\mathbf{K}}\mathbf{T}\}$ if and only if $\mathcal{R}(\boldsymbol{\Gamma}') \subseteq \mathcal{R}(\boldsymbol{\Lambda}')$ and $r[\mathbf{T}\mathbf{X}, \mathbf{T}\boldsymbol{\Sigma}] = r[\mathbf{X}, \boldsymbol{\Sigma}] + r(\mathbf{T}) - n$. In this case, $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} = \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely.

Proof. From (3.15) and (3.25), the difference $\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T}$ can be written as

$$\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T} = \mathbf{Q} + \mathbf{U}_2[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^+]^\perp - \mathbf{U}_1[\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{T}\mathbf{X})^+]^\perp\mathbf{T}, \tag{4.3}$$

where $\mathbf{Q} = [\mathbf{A}\mathbf{X}', \mathbf{B}\boldsymbol{\Sigma}\mathbf{X}^+][\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^+]^\dagger - [\mathbf{A}(\mathbf{T}\mathbf{X})', \mathbf{B}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{T}\mathbf{X})^+][\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{T}\mathbf{X})^+]^\dagger\mathbf{T}$ and $\mathbf{U}_1 \in \mathbb{R}^{s \times m}$ and $\mathbf{U}_2 \in \mathbb{R}^{s \times n}$ are arbitrary. Applying (2.8) to (4.3) gives

$$\min_{\widehat{\mathbf{L}}, \widehat{\mathbf{K}}} r(\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T}) = \min_{\mathbf{U}_1, \mathbf{U}_2} r(\mathbf{Q} + \mathbf{U}_2[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^+]^\perp - \mathbf{U}_1[\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{T}\mathbf{X})^+]^\perp\mathbf{T})$$

$$= r \begin{bmatrix} \mathbf{Q} \\ [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp]^\perp \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \end{bmatrix} - r \begin{bmatrix} [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp]^\perp \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \end{bmatrix}. \quad (4.4)$$

It is easy to obtain by (2.3), (2.4) and elementary block matrix operations (EBMOs) that

$$\begin{aligned} & r \begin{bmatrix} \mathbf{Q} \\ [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp]^\perp \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \end{bmatrix} \\ &= r \begin{bmatrix} \mathbf{Q} & \mathbf{0} & \mathbf{0} \\ \mathbf{I}_n & [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] & \mathbf{0} \\ \mathbf{T} & \mathbf{0} & [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \end{bmatrix} - r[\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] - r[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \\ &= r \begin{bmatrix} \mathbf{0} & -[\mathbf{AX}', \mathbf{B}\boldsymbol{\Sigma X}^\perp] & [\mathbf{A}(\mathbf{TX})', \mathbf{B}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \\ \mathbf{I}_n & [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] & \mathbf{0} \\ \mathbf{T} & \mathbf{0} & [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \end{bmatrix} - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r \begin{bmatrix} \mathbf{0} & -[\mathbf{AX}', \mathbf{B}\boldsymbol{\Sigma X}^\perp] & [\mathbf{A}(\mathbf{TX})', \mathbf{B}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \\ \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{T}[\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] & [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \end{bmatrix} - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r \begin{bmatrix} [\mathbf{AX}', \mathbf{B}\boldsymbol{\Sigma X}^\perp] & [\mathbf{A}(\mathbf{TX})', \mathbf{B}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \\ \mathbf{T}[\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] & [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \end{bmatrix} + n - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r \begin{bmatrix} \mathbf{TXX}' & \mathbf{T}\boldsymbol{\Sigma} & \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}' \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{TX})' \\ \mathbf{AX}' & \mathbf{B}\boldsymbol{\Sigma} & \mathbf{B}\boldsymbol{\Sigma}\mathbf{T}' \end{bmatrix} + n - r(\mathbf{X}) - r(\mathbf{TX}) - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r \begin{bmatrix} \mathbf{TXX}' & \mathbf{T}\boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{TX})' \\ \mathbf{AX}' & \mathbf{B}\boldsymbol{\Sigma} & \mathbf{0} \end{bmatrix} + n - r(\mathbf{X}) - r(\mathbf{TX}) - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r \begin{bmatrix} \mathbf{TXX}' & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \\ \mathbf{AX}' & \mathbf{B}\boldsymbol{\Sigma} \end{bmatrix} + n - r(\mathbf{T}) - r[\mathbf{X}, \boldsymbol{\Sigma}], \quad (4.5) \end{aligned}$$

and

$$\begin{aligned} & r \begin{bmatrix} [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp]^\perp \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \end{bmatrix} \\ &= r \begin{bmatrix} \mathbf{I}_n & [\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] & \mathbf{0} \\ \mathbf{T} & \mathbf{0} & [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \end{bmatrix} - r[\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] - r[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \\ &= r \begin{bmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{T}[\mathbf{XX}', \boldsymbol{\Sigma X}^\perp] & [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp] \end{bmatrix} - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r(\mathbf{T}[\mathbf{XX}', \boldsymbol{\Sigma X}^\perp], [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]) + n - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \end{aligned}$$

$$\begin{aligned}
&= r \begin{bmatrix} \mathbf{TX}\mathbf{X}' & \mathbf{T}\boldsymbol{\Sigma} & \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}' \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{TX})' \end{bmatrix} + n - r(\mathbf{X}) - r(\mathbf{TX}) - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\
&= r \begin{bmatrix} \mathbf{TX}\mathbf{X}' & \mathbf{T}\boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{TX})' \end{bmatrix} + n - r(\mathbf{X}) - r(\mathbf{TX}) - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\
&= r \begin{bmatrix} \mathbf{TX}\mathbf{X}' & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \end{bmatrix} + n - r(\boldsymbol{\Gamma}) - r[\mathbf{X}, \boldsymbol{\Sigma}]. \tag{4.6}
\end{aligned}$$

Substituting (4.5) and (4.6) into (4.4) yields

$$\min_{\widehat{\mathbf{L}}, \widehat{\mathbf{K}}} r(\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T}) = r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} - r(\boldsymbol{\Lambda}). \tag{4.7}$$

Setting the right-hand side of (4.7) equal to zero and applying Lemma 2.2(b) yields the equivalent condition in (a). Applying (2.8) to (4.3) yields

$$\begin{aligned}
\min_{\widehat{\mathbf{L}}} r(\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T}) &= \min_{\mathbf{U}_2} r(\mathbf{Q} + \mathbf{U}_2[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp - \mathbf{U}_1[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T}) \\
&= r \begin{bmatrix} \mathbf{Q} - \mathbf{U}_1[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp \end{bmatrix} - r([\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp) \\
&= r \begin{bmatrix} \mathbf{Q} - \mathbf{U}_1[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp \end{bmatrix} - n + r[\mathbf{X}, \boldsymbol{\Sigma}] \tag{4.8}
\end{aligned}$$

and by (2.9) and (4.5),

$$\begin{aligned}
&\max_{\mathbf{U}_1} r \begin{bmatrix} \mathbf{Q} - \mathbf{U}_1[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp \end{bmatrix} \\
&= \max_{\mathbf{U}_1} r \left(\begin{bmatrix} \mathbf{Q} \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp \end{bmatrix} - \begin{bmatrix} \mathbf{I}_s \\ \mathbf{0} \end{bmatrix} \mathbf{U}_1[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \right) \\
&= \min \left\{ r \begin{bmatrix} \mathbf{Q} \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \end{bmatrix}, r \begin{bmatrix} \mathbf{Q} & \mathbf{I}_s \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]^\perp & \mathbf{0} \end{bmatrix} \right\} \\
&= \min \left\{ r \begin{bmatrix} \mathbf{Q} \\ [\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp] \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T} \end{bmatrix}, s + n - r[\mathbf{X}, \boldsymbol{\Sigma}] \right\} \\
&= \min \left\{ r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} + n - r(\mathbf{X}) - r[\mathbf{X}, \boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}], s + n - r[\mathbf{X}, \boldsymbol{\Sigma}] \right\} \\
&= \min \left\{ s, r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} - r(\boldsymbol{\Gamma}) \right\} + n - r[\mathbf{X}, \boldsymbol{\Sigma}]. \tag{4.9}
\end{aligned}$$

Combining (4.8) and (4.9) yields

$$\max_{\widehat{\mathbf{K}}} \min_{\widehat{\mathbf{L}}} r(\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T}) = \min \left\{ s, r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} - r(\boldsymbol{\Gamma}) \right\}. \tag{4.10}$$

Setting the right-hand side of (4.10) equal to zero yields $r\begin{bmatrix} \Lambda \\ \Gamma \end{bmatrix} = r(\Gamma)$. Thus, the statement in (b) holds.

By a similar approach, we can obtain

$$\max_{\widehat{\mathbf{L}}} \min_{\widehat{\mathbf{K}}} r(\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T}) = \min \left\{ s, r\begin{bmatrix} \Lambda \\ \Gamma \end{bmatrix} + n - r(\mathbf{T}) - r(\mathbf{X}) - r[\mathbf{X}, \Sigma] \right\}, \quad (4.11)$$

as required for the statement in (c). Combining (b) and (c) yields (d). \square

Theorem 4.2. Let $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ and $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ be as given in (3.15) and (3.25), respectively, and let Λ and Γ be as given in (4.2). Then, the following six statements are equivalent:

- (a) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds definitely.
- (b) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds with probability 1.
- (c) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \supseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.
- (d) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \subseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.
- (e) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} = \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.
- (f) $\mathcal{R}(\Gamma') \subseteq \mathcal{R}(\Lambda')$.

Proof. It can be seen from Lemma 2.6 and Definition 2.8(c) that (a) is equivalent to

$$\min_{\widehat{\mathbf{L}}, \widehat{\mathbf{K}}} r((\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T})[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}]) = 0. \quad (4.12)$$

Substituting the coefficient matrices in (3.15) and (3.25) into (4.12) and simplifying, we obtain

$$\mathbf{U}_1[\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}]^{\perp}\mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}] = \mathbf{J},$$

where $\mathbf{J} = [\mathbf{A}\mathbf{X}', \mathbf{B}\Sigma\mathbf{X}^{\perp}] - [\mathbf{A}(\mathbf{T}\mathbf{X})', \mathbf{B}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}][\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}]^{\perp}\mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}]$ and $\mathbf{U}_1 \in \mathbb{R}^{s \times m}$ is arbitrary. From Lemma 2.5, the matrix equation is solvable for \mathbf{U}_1 if and only if

$$r\begin{bmatrix} \mathbf{J} \\ [\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}]^{\perp}\mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}] \end{bmatrix} = r([\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}]^{\perp}\mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}]). \quad (4.13)$$

Applying (2.3) and (2.4), and simplifying, leads to

$$\begin{aligned} & r\begin{bmatrix} \mathbf{J} \\ [\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}]^{\perp}\mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}] \end{bmatrix} \\ &= r\begin{bmatrix} \mathbf{J} & \mathbf{0} \\ \mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}] & [\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}] \end{bmatrix} - r[\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}] \\ &= r\begin{bmatrix} [\mathbf{A}\mathbf{X}', \mathbf{B}\Sigma\mathbf{X}^{\perp}] & [\mathbf{A}(\mathbf{T}\mathbf{X})', \mathbf{B}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}] \\ \mathbf{T}[\mathbf{X}\mathbf{X}', \Sigma\mathbf{X}^{\perp}] & [\mathbf{T}\mathbf{X}(\mathbf{T}\mathbf{X})', \mathbf{T}\Sigma\mathbf{T}'(\mathbf{T}\mathbf{X})^{\perp}] \end{bmatrix} - r[\mathbf{T}\mathbf{X}, \mathbf{T}\Sigma] \\ &= r\begin{bmatrix} \mathbf{T}\mathbf{X}\mathbf{X}' & \mathbf{T}\Sigma & \mathbf{T}\Sigma\mathbf{T}' \\ \mathbf{0} & \mathbf{X}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & (\mathbf{T}\mathbf{X})' \\ \mathbf{A}\mathbf{X}' & \mathbf{B}\Sigma & \mathbf{B}\Sigma\mathbf{T}' \end{bmatrix} - r(\mathbf{X}) - r(\mathbf{T}\mathbf{X}) - r[\mathbf{T}\mathbf{X}, \mathbf{T}\Sigma] \end{aligned}$$

$$= r \begin{bmatrix} \mathbf{TX}\mathbf{X}' & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \\ \mathbf{A}\mathbf{X}' & \mathbf{B}\boldsymbol{\Sigma} \end{bmatrix} - r(\mathbf{X}) - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] = r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} - r(\boldsymbol{\Gamma}), \quad (4.14)$$

and

$$\begin{aligned} & r([\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T}[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) \\ &= r([\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp], \mathbf{T}[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] \\ &= r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] - r[\mathbf{TX}, \mathbf{T}\boldsymbol{\Sigma}] = 0. \end{aligned} \quad (4.15)$$

Substituting (4.14) and (4.15) into (4.13) leads to $r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} = r(\boldsymbol{\Gamma})$, thus establishing the equivalence of (a) and (e).

From Lemma 2.6 and Definition 2.8(d) that (b) is equivalent to

$$\max_{\widehat{\mathbf{K}}} \min_{\widehat{\mathbf{L}}} r((\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T})[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) = 0. \quad (4.16)$$

From (2.7), (3.15), (3.23), (3.25) and (4.14),

$$\begin{aligned} & \max_{\widehat{\mathbf{K}}} \min_{\widehat{\mathbf{L}}} r((\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T})[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) \\ &= \max_{\widehat{\mathbf{K}}} r(\mathbf{J} - \mathbf{U}_1[\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T}[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) \\ &= \min \left\{ r \begin{bmatrix} \mathbf{J} \\ [\mathbf{TX}(\mathbf{TX})', \mathbf{T}\boldsymbol{\Sigma}\mathbf{T}'(\mathbf{TX})^\perp]^\perp \mathbf{T}[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp] \end{bmatrix}, s \right\}. \end{aligned} \quad (4.17)$$

Setting the right-hand side of (4.17) equal to zero yields $r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} = r(\boldsymbol{\Gamma})$, thus establishing the equivalence of (b) and (e).

Similarly, we are able to obtain

$$\max_{\widehat{\mathbf{L}}} \min_{\widehat{\mathbf{K}}} r((\widehat{\mathbf{L}} - \widehat{\mathbf{K}}\mathbf{T})[\mathbf{X}\mathbf{X}', \boldsymbol{\Sigma}\mathbf{X}^\perp]) = r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} - r(\boldsymbol{\Gamma}). \quad (4.18)$$

Thus, (4.18) is equivalent to $r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} = r(\boldsymbol{\Gamma})$. Combining results (b) and (c) leads to the equivalence of (d) and (e). \square

Combining Theorems 4.1 and 4.2, we obtain the following result:

Corollary 4.3. *Let $\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})$ and $\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})$ be as given in (3.15) and (3.25), respectively, and let $\boldsymbol{\Lambda}$ and $\boldsymbol{\Gamma}$ be as given in (4.2). Then, the following six statistical statements are equivalent:*

- (a) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds definitely.
- (b) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \supseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds definitely.
- (c) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \cap \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\} \neq \emptyset$ holds with probability 1.
- (d) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \supseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.

- (e) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} \subseteq \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.
 (f) $\{\text{BLMBP}_{\mathcal{M}}(\boldsymbol{\tau})\} = \{\text{BLMBP}_{\mathcal{N}}(\boldsymbol{\tau})\}$ holds with probability 1.

The results in the above theorems and corollaries can be simplified further for different choices of the matrices in (1.3), such as, $\mathbf{A} = \mathbf{K}$ and $\mathbf{B} = \mathbf{0}$. Hence, many specified conclusions can further be obtained on the relationships between a GLM and its TGLMs. Assume that $\boldsymbol{\tau}$ in (1.3) is predictable under (1.1) and (1.5). We obtain that the BLMBP of $\boldsymbol{\tau}$ is just its BLUP, and the main results in the paper are the classic theory on the BLUPs of $\boldsymbol{\tau}$ under (1.1) and (1.5). Therefore, these works are certain extensions of the classic BLUP theory.

5. An example

Assume that a concrete form of \mathcal{M} in (1.1) is given by

$$\mathcal{M} : \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix}, \quad \mathbb{E} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \mathbf{0}, \quad \text{Cov} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{bmatrix} = \sigma^2 \begin{bmatrix} \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_2} \end{bmatrix}.$$

In this case, taking two transformation matrices $\mathbf{T}_1 = [\mathbf{I}_{n_1}, \mathbf{0}]$ and $\mathbf{T}_2 = [\mathbf{0}, \mathbf{I}_{n_2}]$, we obtain the following two sub-sample models:

$$\begin{aligned} \mathcal{M}_1 : \mathbf{y}_1 &= \mathbf{X}_1 \boldsymbol{\theta}_1 + \mathbf{v}_1, \quad \mathbb{E}(\mathbf{v}_1) = \mathbf{0}, \quad \text{Cov}(\mathbf{v}_1) = \sigma^2 \mathbf{I}_{n_1}, \\ \mathcal{M}_2 : \mathbf{y}_2 &= \mathbf{X}_2 \boldsymbol{\theta}_2 + \mathbf{v}_2, \quad \mathbb{E}(\mathbf{v}_2) = \mathbf{0}, \quad \text{Cov}(\mathbf{v}_2) = \sigma^2 \mathbf{I}_{n_2}, \end{aligned}$$

where it is assumed that $\mathbf{y}_i \in \mathbb{R}^{n_i \times 1}$, $\mathbf{X}_i \in \mathbb{R}^{n_i \times p_i}$, $\boldsymbol{\theta}_i \in \mathbb{R}^{p_i \times 1}$, $\mathbf{v}_i \in \mathbb{R}^{n_i \times 1}$, $n = n_1 + n_2$, $p = p_1 + p_2$. For illustrating the results in Section 4, let $\mathbf{A} = [\mathbf{X}_1, \mathbf{0}]$, $\mathbf{B} = \mathbf{0}$ and $\mathbf{A} = [\mathbf{0}, \mathbf{X}_2]$, $\mathbf{B} = \mathbf{0}$ in (1.3), respectively. From Theorems 4.1 and 4.2,

$$r \begin{bmatrix} \boldsymbol{\Lambda} \\ \boldsymbol{\Gamma} \end{bmatrix} = r \begin{bmatrix} \mathbf{TXX}' & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \\ \mathbf{AX}' & \mathbf{B}\boldsymbol{\Sigma} \end{bmatrix} = r \begin{bmatrix} \mathbf{X}_1 \mathbf{X}_1' & \sigma^2 \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_2' \\ \mathbf{X}_1 \mathbf{X}_1' & \mathbf{0} & \mathbf{0} \end{bmatrix} = r \begin{bmatrix} \mathbf{0} & \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_2' \\ \mathbf{X}_1 \mathbf{X}_1' & \mathbf{0} & \mathbf{0} \end{bmatrix} = n_1 + r(\mathbf{X}),$$

$$r(\boldsymbol{\Lambda}) = r \begin{bmatrix} \mathbf{TXX}' & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \end{bmatrix} = r \begin{bmatrix} \mathbf{TX} & \mathbf{T}\boldsymbol{\Sigma} \\ \mathbf{0} & \mathbf{X}' \end{bmatrix} = r \begin{bmatrix} \mathbf{X}_1 & \sigma^2 \mathbf{I}_{n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_1' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_2' \end{bmatrix} = n_1 + r(\mathbf{X}),$$

$$r(\mathbf{T}) + r(\mathbf{X}) + r[\mathbf{X}, \boldsymbol{\Sigma}] - n = n_1 + r(\mathbf{X}).$$

Obviously, the equivalent conditions all hold in Theorems 4.1 and 4.2. Thus, we can easily describe the relations between the corresponding estimators.

6. Conclusions

We have provided algebraic and statistical analysis of a biased prediction problem when a joint parametric vector is unpredictable under a given GLM, and obtained an abundance of exact formulas and facts about the BLMBPs of the joint parametric vector in the contexts of a GLM and its TGLMs. All the findings in this article are technically formulated or denoted in certain analytical expressions or

explicit assertions through the surprise use of specified matrix analysis tools and techniques. Hence, it is not difficult to understand these results and facts from both mathematical and statistical aspects. In view of this fact, we can take these obtained in the preceding sections as a group of theoretical contributions in the statistical inference under general linear model assumptions. Consequently, we are able to utilize the statistical methods developed in this article to provide additional insight into various concrete inference problems and subjects related to GLMs. Correspondingly, we point out that the main conclusions presented in this work have certain significant applications in the field of inverse scattering problems. The reader is referred to [8, 31, 32] on the topic of inverse scattering problems.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

We are grateful to anonymous reviewers for their helpful comments and suggestions. The second author was supported in part by the Shandong Provincial Natural Science Foundation #ZR2019MA065.

Conflict of Interest

The authors declare that they have no conflicts of interest.

References

1. J. K. Baksalary, R. Kala, Linear transformations preserving best linear unbiased estimators in a general Gauss-Markoff model, *Ann. Stat.*, **9** (1981), 913–916.
2. B. Dong, W. Guo, Y. Tian, On relations between BLUEs under two transformed linear models, *J. Multivariate Anal.*, **131** (2014), 279–292. <https://doi.org/10.1016/j.jmva.2014.07.005>
3. H. Drygas, Sufficiency and completeness in the general Gauss-Markov model, *Sankhyā A*, **45** (1983), 88–98.
4. N. Güler, On relations between BLUPs under two transformed linear random-effects models, *Commun. Stat. Simul. Comput.*, **51** (2022), 5099–5125. <https://doi.org/10.1080/03610918.2020.1757709>
5. N. Güler, M. E. Büyükkaya, Notes on comparison of covariance matrices of BLUPs under linear random-effects model with its two subsample models, *Iran. J. Sci. Tech. Trans. A*, **43** (2019), 2993–3002. <https://doi.org/10.1007/s40995-019-00785-3>
6. N. Güler, M. E. Büyükkaya, Inertia and rank approach in transformed linear mixed models for comparison of BLUPs, *Commun. Stat. Theor. Meth.*, **52** (2023), 3108–3123. <https://doi.org/10.1080/03610926.2021.1967397>
7. S. J. Haslett, J. Isotalo, Y. Liu, S. Puntanen, Equalities between OLSE, BLUE and BLUP in the linear model, *Stat. Papers*, **55** (2014), 543–561. <https://doi.org/10.1007/s00362-013-0500-7>

8. Y. He, H. Liu, X. Wang, A novel quantitative inverse scattering scheme using interior resonant modes, *Inverse Probl.*, **39** (2023), 085002. <https://doi.org/10.1088/1361-6420/acdc49>
9. R. Kala, P. R. Pordzik, Estimation in singular partitioned, reduced or transformed linear models, *Stat. Papers*, **50** (2009), 633–638. <https://doi.org/10.1007/s00362-007-0097-9>
10. 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>
11. A. Markiewicz, S. Puntanen, Further properties of linear prediction sufficiency and the BLUPs in the linear model with new observations, *Afrika Stat.*, **13** (2018), 1511–1530. <https://doi.org/10.16929/as/1511.117>
12. G. Marsaglia, G. P. H. Styan, Equalities and inequalities for ranks of matrices, *Linear Multilinear Algebra*, **2** (1974), 269–292. <https://doi.org/10.1080/03081087408817070>
13. S. K. Mitra, Generalized inverse of matrices and applications to linear models, In: *Handbook of Statistics*, **1** (1980), 471–512. [https://doi.org/10.1016/S0169-7161\(80\)80045-9](https://doi.org/10.1016/S0169-7161(80)80045-9)
14. C. H. Morrell, J. D. Pearson, L. J. Brant, Linear transformations of linear mixed-effects models, *Amer. Stat.*, **51** (1997), 338–343. <https://doi.org/10.1080/00031305.1997.10474409>
15. R. Penrose, A generalized inverse for matrices, *Proc. Cambridge Phil. Soc.*, **51** (1955), 406–413. <https://doi.org/10.1017/S0305004100030401>
16. S. Puntanen, G. P. H. Styan, J. Isotalo, *Matrix Tricks for Linear Statistical Models: Our Personal Top Twenty*, Berlin: Springer, 2011.
17. C. R. Rao, *Linear Statistical Inference and Its Applications*, New York: Wiley, 1973.
18. C. R. Rao, Choice of best linear estimators in the Gauss-Markoff model with a singular dispersion matrix, *Commun. Stat. Theor. Meth.*, **7** (1978), 1199–1208.
19. C. R. Rao, S. K. Mitra, *Generalized Inverse of Matrices and Its Applications*, New York: Wiley, 1971.
20. J. Shao, J. Zhang, A transformation approach in linear mixed-effects models with informative missing responses, *Biometrika*, **102** (2015), 107–119. <https://doi.org/10.1093/biomet/asu069>
21. Y. Tian, The maximal and minimal ranks of some expressions of generalized inverses of matrices, *SEA Bull. Math.*, **25** (2002), 745–755. <https://doi.org/10.1007/s100120200015>
22. Y. Tian, More on maximal and minimal ranks of Schur complements with applications, *Appl. Math. Comput.*, **152** (2004), 675–692. [https://doi.org/10.1016/S0096-3003\(03\)00585-X](https://doi.org/10.1016/S0096-3003(03)00585-X)
23. Y. Tian, On properties of BLUEs under general linear regression models, *J. Stat. Plann. Inference*, **143** (2013), 771–782. <https://doi.org/10.1016/j.jspi.2012.10.005>
24. Y. Tian, A new derivation of BLUPs under random-effects model, *Metrika*, **78** (2015), 905–918. <https://doi.org/10.1007/s00184-015-0533-0>
25. Y. Tian, Transformation approaches of linear random-effects models, *Stat. Meth. Appl.*, **26** (2017), 583–608. <https://doi.org/10.1007/s10260-017-0381-3>
26. Y. Tian, Matrix rank and inertia formulas in the analysis of general linear models, *Open Math.*, **15** (2017), 126–150. <https://doi.org/10.1515/math-2017-0013>

27. Y. Tian, S. Cheng, The maximal and minimal ranks of A-BXC with applications, *New York J. Math.*, **9** (2003), 345–362.
28. Y. Tian, B. Jiang, A new analysis of the relationships between a general linear model and its mis-specified forms, *J. Korean Stat. Soc.*, **46** (2017), 182–193. <https://doi.org/10.1016/j.jkss.2016.08.004>
29. Y. Tian, S. Puntanen, On the equivalence of estimations under a general linear model and its transformed models, *Linear Algebra Appl.*, **430** (2009), 2622–2641. <https://doi.org/10.1016/j.laa.2008.09.016>
30. C. Xie, Linear transformations preserving best linear minimum bias linear estimators in a Gauss-Markoff model, *Appl. Math. J. Chin. Univer. Ser. A*, **9** (1994), 429–434.
31. W. Yin, W. Yang, H. Liu, A neural network scheme for recovering scattering obstacles with limited phaseless far-field data, *J. Comput. Phys.*, **417** (2020), 109594. <https://doi.org/10.1016/j.jcp.2020.109594>
32. Y. Yin, W. Yin, P. Meng, H. Liu, The interior inverse scattering problem for a two-layered cavity using the Bayesian method, *Inverse Probl. Imaging*, **16** (2022), 673–690. <http://dx.doi.org/10.3934/ipi.2021069>
33. B. Zhang, The BLUE and MINQUE in Gauss-Markoff model with linear transformation of the observable variables, *Acta Math. Sci.*, **27** (2007), 203–210. [https://doi.org/10.1016/S0252-9602\(07\)60018-6](https://doi.org/10.1016/S0252-9602(07)60018-6)



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)