Statistics and Probability Letters 112 (2016) 105–112

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On connections among OLSEs and BLUEs of whole and partial parameters under a general linear model

Yongge Tian [a,*], Xuan Zhang [b]

[a] *China Economics and Management Academy, Central University of Finance and Economics, Beijing, China*

[b] *School of Mathematics and Statistics, Zhongnan University of Economics and Law, Wuhan, China*

ARTICLE INFO

Article history:
Received 13 October 2015
Received in revised form 26 January 2016
Accepted 26 January 2016
Available online 3 February 2016

MSC:
62H12
62J05

Keywords:
Partitioned linear model
Partial parameter vector
OLSE
BLUE

ABSTRACT

This paper presents a new investigation to the connections among the ordinary least squares estimators (OLSEs) and the best linear unbiased estimators (BLUEs) of the whole and partial mean parameter vectors in a multiple partitioned linear model. We first give some general results on the equivalence of the OLSEs and the BLUEs under a general linear model, and derive some new facts on the connections among the OLSEs and the BLUEs of the whole and partial mean parameter vectors in the model.

1. Introduction

Linear statistical models were the first type of models to be studied rigorously in statistical analysis, which were regarded without doubt as a noble and magnificent part in current statistical theory. Recall that the ordinary least squares estimator (OLSE for short) and the best linear unbiased estimator (BLUE for short) of unknown parameters in a general linear model are two fundamental and useful types of estimator in statistical analysis. These two types of estimator are defined from two optimality criteria, while the expressions of OLSEs and BLUEs can be derived from certain algebraic operations of the observed response vectors, the given model matrices, and the covariance matrices of the error terms in the model. Based on the exact algebraic expressions of OLSEs and BLUEs under linear models, as well as various algebraic tools in matrix theory, people derived many simple and useful properties of OLSEs and BLUEs in the literature, and established a systematic theory on OLSEs and BLUEs and their applications. Even so, there are many new problems on OLSEs and BLUEs that can be proposed and studied from theoretical and applied points of view. Although these estimators have different performances in the statistical inference of general linear models, people were really interested in the connections between the OLSE and the BLUE, and established many identifying conditions for the OLSE and the BLUE to be equivalent. In fact, general linear models are only type of statistical models that have complete and solid supports from the analytical tools in linear algebra and matrix theory; see, e.g., Puntanen et al. (2011). It is just based on this fact that linear statistical models attract a few of linear algebraists to consider applications of their matrix contributions in statistical analysis. On the other hand, approaches of the linear statistical models also bring many theoretical considerations in matrix analysis.

* Corresponding author.
E-mail addresses: yongge.tian@gmail.com (Y. Tian), zhangx@znufe.edu.cn (X. Zhang).

http://dx.doi.org/10.1016/j.spl.2016.01.019
0167-7152/

Linear models can be presented in certain partitioned forms, which are usually used in the estimations of partial unknown parameters in the models as well as in the investigations of certain reduced models or small models associated with the original models. Let us consider a general partitioned linear model defined by

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

where $\mathbf{y} \in \mathbb{R}^{n\times 1}$ is an observable random vector, $\mathbf{X}_i \in \mathbb{R}^{n\times p_i}$ is a known matrix of arbitrary rank with $\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_k]$, $\boldsymbol{\beta}_i \in \mathbb{R}^{p_i\times 1}$ are fixed but unknown parameter vector with $\boldsymbol{\beta} = [\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_k']'$ and $p = p_1 + \cdots + p_k$, $i = 1, \ldots, k$, $E(\boldsymbol{\varepsilon})$ and $D(\boldsymbol{\varepsilon})$ denote the expectation vector and the dispersion matrix of the random error vector $\boldsymbol{\varepsilon} \in \mathbb{R}^{n\times 1}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ is a known nonnegative definite matrix of arbitrary rank, and σ^2 is an unknown positive parameter.

Let

$$\mathbf{Y}_i = [\mathbf{0}, \ldots, \mathbf{X}_i, \ldots, \mathbf{0}], \qquad \mathbf{Z}_i = [\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}, \mathbf{0}, \mathbf{X}_{i+1}, \ldots, \mathbf{X}_k], \quad i = 1, \ldots, k. \tag{1.2}$$

Then, the model matrix $\mathbf{X}$ in (1.1) can be decomposed as

$$\mathbf{X} = \mathbf{Y}_i + \mathbf{Z}_i = \mathbf{Y}_1 + \cdots + \mathbf{Y}_k, \quad i = 1, \ldots, k. \tag{1.3}$$

Correspondingly, the partial mean parameter vectors $\mathbf{X}_i\boldsymbol{\beta}_i$ on the right-hand side of (1.1) can be rewritten as

$$\mathbf{X}_i\boldsymbol{\beta}_i = \mathbf{Y}_i\boldsymbol{\beta}, \quad i = 1, \ldots, k. \tag{1.4}$$

Throughout this paper, $\mathbb{R}^{m\times n}$ stands for the collection of all $m \times n$ real matrices. The symbols $\mathbf{A}'$, $r(\mathbf{A})$ and $\mathscr{R}(\mathbf{A})$ stand for the transpose, the rank and the range (column space) of a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, respectively. $\mathbf{I}_m$ denotes the identity matrix of order m. The Moore–Penrose inverse of $\mathbf{A} \in \mathbb{R}^{m\times n}$, denoted by $\mathbf{A}^+$, is defined to be the unique solution $\mathbf{G}$ to the four matrix equations $\mathbf{AGA} = \mathbf{A}$, $\mathbf{GAG} = \mathbf{G}$, $(\mathbf{AG})' = \mathbf{AG}$ and $(\mathbf{GA})' = \mathbf{GA}$. Further, denote $\mathbf{P}_\mathbf{A} = \mathbf{AA}^+$, $\mathbf{A}^\perp = \mathbf{E}_\mathbf{A} = \mathbf{I}_m - \mathbf{AA}^+$ and $\mathbf{F}_\mathbf{A} = \mathbf{I}_n - \mathbf{A}^+\mathbf{A}$. 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. It is well known that the Löwner partial ordering is a surprisingly strong and useful property between two symmetric matrices. For more issues about the Löwner partial ordering of symmetric matrices and applications in statistic analysis, see, e.g., Puntanen et al. (2011).

It is well known that methods for estimating vectors of the unknown parameters in general linear models often make use of inverses or generalized inverses of matrices in their theoretical development; see, e.g., Aitken (1934), Rao (1973) and Zyskind and Martin (1969). In fact, the exact algebraic expressions of the OLSEs and BLUEs of $\mathbf{X}\boldsymbol{\beta}$ and $\mathbf{X}_i\boldsymbol{\beta}_i$ can all be derived from routine operations of the given matrices in the model and their generalized inverses; see, e.g., Drygas (1970), Graybill (1961), Rao (1973) and Searle (1971), while many properties of the OLSEs and BLUEs can be derived from the formulas. At the same time, people are also interested in the connections between the OLSE and BLUE, and like to know when the OLSE is the BLUE. Many identifying conditions for the OLSE to be the BLUE under a general linear model were established. Some backgrounds on the equivalence of OLSEs and BLUEs under general linear models can be found, e.g., in Alalouf and Styan (1984), Baksalary and Kala (1977), Baksalary and Puntanen (1990), Baksalary et al. (1990), Haberman (1974), Haslett et al. (2014), Haslett and Puntanen (2010a,b), Haslett and Puntanen (2011), Herzberg and Aleong (1995), Isotalo and Puntanen (2009), Kruskal (1968), McElroy (1967), Puntanen and Styan (1989) and Styan (1973).

We reconsider in this paper the relationships among the OLSEs and BLUEs of the whole and partial mean parameter vectors in (1.1). We first present some known necessary and sufficient conditions for $\mathrm{OLSE}(\mathbf{K}\boldsymbol{\beta}) = \mathrm{BLUE}(\mathbf{K}\boldsymbol{\beta})$ and $\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta})$ to hold. We then prove the following two equivalent statements for the OLSEs and BLUEs of the whole and partial mean parameter vectors in (1.1):

$$\mathrm{OLSE}(\mathbf{X}\boldsymbol{\beta}) = \mathrm{BLUE}(\mathbf{X}\boldsymbol{\beta}) \Longleftrightarrow \mathrm{OLSE}(\mathbf{X}_i\boldsymbol{\beta}_i) = \mathrm{BLUE}(\mathbf{X}_i\boldsymbol{\beta}_i), \quad i = 1, \ldots, k. \tag{1.5}$$

The equivalence in (1.5) for $k = 2$ was shown in Tian and Zhang (2011). It should be pointed out that these implications cannot be directly obtained from the definitions of OLSEs and BLUEs, or from the algebraic expressions of these OLSEs and BLUEs under (1.1).

Statistical methods in many areas of application often involve mathematical computations with vectors and matrices. In particular, formulas and algebraic tricks for handling matrices in linear algebra and matrix theory play important roles in the derivation of estimators and characterizations of their properties under linear regression models. Recall that the rank of matrix is a conceptual foundation in linear algebra and matrix theory, which is the most significant finite nonnegative integer in reflecting intrinsic properties of matrices. It has long history to establish rank formulas for block matrices and use the formulas in statistical inferences, and a pioneer work in this aspect can be found in Guttman (1944). The intriguing connections between generalized inverses of matrices and rank formulas of matrices were recognized in 1970s, and a seminal work on rank formulas for matrices and their generalized inverses was presented in Marsaglia and Styan (1974). In order to establish and characterize various possible equalities for estimators under linear statistical models, and to simplify various matrix equalities composed by the Moore–Penrose inverses of matrices, we need the following well-known rank formulas for matrices and their Moore–Penrose generalized inverses to make the paper self-contained.