Journal of Multivariate Analysis 162 (2017) 193–214

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

On additive decompositions of estimators under a multivariate general linear model and its two submodels

Bo Jiang [a], Yongge Tian [b,*]

[a] College of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai, China
[b] China Economics and Management Academy, Central University of Finance and Economics, Beijing, China

ARTICLE INFO

Article history:
Received 8 December 2016
Available online 29 September 2017

Keywords:
BLUE
Decomposition equality
Estimability
MGLM
OLSE
Parameter matrix
Submodel

ABSTRACT

Parameters from linear regression models are often estimated by the ordinary least squares estimator (OLSE) or by the best linear unbiased estimator (BLUE). These estimators can be written in analytical form, so that it is not difficult to describe their performances under various model assumptions. In this paper, we study the problem of additive decompositions of OLSEs and BLUEs of parameter spaces in a full multivariate general linear model (MGLM) and in two specific submodels. We establish necessary and sufficient conditions for the validity of various identities involving the OLSEs and BLUEs of whole and partial mean parameter matrices under the MGLM and two smaller MGLMs.

1. Introduction

In linear regression modeling, it is usual to decompose the full regressors as a sum of partial regressors in order to identify those that are the most important or to rank them. Such a decomposition makes it possible to determine the roles of the partial regressors, and to derive estimators of partial unknown parameters under such a partitioned linear model. Reduced models are usually associated with a partitioned linear model, and it is of interest to establish additive decompositions of the corresponding estimators.

Let us consider a multivariate general linear model (MGLM) defined by

$$\mathcal{M}: \begin{cases} \mathbf{Y} = \mathbf{X}\Theta + \Psi = \mathbf{X}_1\Theta_1 + \mathbf{X}_2\Theta_2 + \Psi, \\ \mathrm{E}(\Psi) = \mathbf{0}, \quad \mathrm{D}(\overrightarrow{\Psi}) = \mathrm{cov}(\overrightarrow{\Psi}, \overrightarrow{\Psi}) = \sigma^2(\mathbf{\Sigma}_2 \otimes \mathbf{\Sigma}_1), \end{cases} \tag{1}$$

where $\mathbf{Y} \in \mathbb{R}^{n\times m}$ is a matrix of observable dependent variables which comes from an experimental design giving rise to n observations, and $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2) \in \mathbb{R}^{n\times p}$ is the model matrix of arbitrary rank, and for $i \in \{1, 2\}$, $\mathbf{X}_i \in \mathbb{R}^{n\times p_i}$. In this setup, $\Theta = (\Theta_1^\top, \Theta_2^\top)^\top \in \mathbb{R}^{p\times m}$ is matrix of fixed but unknown parameters, and for $i \in \{1, 2\}$, $\Theta_i \in \mathbb{R}^{p_i\times m}$ with $p = p_1 + p_2$. Furthermore, $\Psi \in \mathbb{R}^{n\times m}$ is a matrix of randomly distributed error terms with zero mean matrix, E and D denote expectation and dispersion matrix, respectively, and $\mathbf{\Sigma}_1 = (\sigma_{1ij}) \in \mathbb{R}^{n\times n}$ and $\mathbf{\Sigma}_2 = (\sigma_{2ij}) \in \mathbb{R}^{m\times m}$ are two known nonnegative definite matrices of arbitrary rank while σ^2 is an arbitrary positive scaling factor. Eq. (1) is also called a general multivariate Gauss–Markov model in the statistical literature.

MGLMs are fairly straightforward extensions of univariate general linear models (UGLMs), in which several response variables are regressed on a given set of explanatory variables. Such models are useful in that they provide a more complete picture of a global problem involving a number of dependent and independent variables. They occur, e.g., in analysis of

* Corresponding author.
E-mail addresses: jiangbo@email.cufe.edu.cn (B. Jiang), yonggetian@cufe.edu.cn (Y. Tian).

https://doi.org/10.1016/j.jmva.2017.09.007
0047-259X/

variance (ANOVA), analysis of covariance (ANCOVA), multivariate analysis of variance (MANOVA), analysis of repeated measurements, factor analysis models, as well as in many areas of applications.

It is common in regression analysis to rewrite a MGLM in partitioned form for inference purposes. Often, we also meet with transformations of a model, and use these transformations to estimate and predict the parameter space in the model. Linear transformations are among the simplest; they consist of pre-multiplying both sides of (1) by two matrices $\mathbf{X}_1^{\perp}$ and $\mathbf{X}_2^{\perp}$ to yield the following pair of transformed MGLMs

$$\mathcal{M}_1 : \mathbf{X}_2^{\perp}\mathbf{Y} = \mathbf{X}_2^{\perp}\mathbf{X}_1\Theta_1 + \mathbf{X}_2^{\perp}\Psi, \quad \mathrm{E}(\mathbf{X}_2^{\perp}\Psi) = \mathbf{0}, \quad \mathrm{D}(\overrightarrow{\mathbf{X}_2^{\perp}\Psi}) = \sigma^2(\mathbf{\Sigma}_2 \otimes \mathbf{X}_2^{\perp}\mathbf{\Sigma}_1\mathbf{X}_2^{\perp}), \tag{2}$$

$$\mathcal{M}_2 : \mathbf{X}_1^{\perp}\mathbf{Y} = \mathbf{X}_1^{\perp}\mathbf{X}_2\Theta_2 + \mathbf{X}_1^{\perp}\Psi, \quad \mathrm{E}(\mathbf{X}_1^{\perp}\Psi) = \mathbf{0}, \quad \mathrm{D}(\overrightarrow{\mathbf{X}_1^{\perp}\Psi}) = \sigma^2(\mathbf{\Sigma}_2 \otimes \mathbf{X}_1^{\perp}\mathbf{\Sigma}_1\mathbf{X}_1^{\perp}), \tag{3}$$

which are usually called correctly-reduced versions of $\mathcal{M}$ in (1); see [10,12] for the corresponding expositions. Since the two linear transformations $\mathbf{X}_2^{\perp}\mathbf{Y}$ and $\mathbf{X}_1^{\perp}\mathbf{Y}$ are singular, each of the model equations in $\mathcal{M}_1$ and $\mathcal{M}_2$ is not equivalent to the full model equation in $\mathcal{M}$. An advantage of formulating $\mathcal{M}_1$ and $\mathcal{M}_2$ arises from the fact that the partial parameter matrices Θ_1 and Θ_2 in $\mathcal{M}$ do not occur in $\mathcal{M}_2$ and $\mathcal{M}_1$, respectively. Thus we can estimate Θ_1 and Θ_2 in $\mathcal{M}$ individually from $\mathcal{M}_1$ and $\mathcal{M}_2$.

In many situations, we have to work with these derived models, hoping that inference results obtained from $\mathcal{M}_1$ and $\mathcal{M}_2$ are equivalent to those that would correspond to $\mathcal{M}$. The search for connections between estimators under original models and their reduced models is called a linear sufficiency problem in statistical inference, which was first introduced in [2,4] and was considered in the statistical literature.

Two other submodels associated with $\mathcal{M}$ in (1) are given by

$$\mathcal{N}_1 : \mathbf{Y} = \mathbf{X}_1\Theta_1 + \Psi_1, \quad \mathrm{E}(\Psi_1) = \mathbf{0}, \quad \mathrm{D}(\overrightarrow{\Psi}_1) = \sigma^2(\mathbf{\Sigma}_2 \otimes \mathbf{\Sigma}_1), \tag{4}$$

$$\mathcal{N}_2 : \mathbf{Y} = \mathbf{X}_2\Theta_2 + \Psi_2, \quad \mathrm{E}(\Psi_2) = \mathbf{0}, \quad \mathrm{D}(\overrightarrow{\Psi}_2) = \sigma^2(\mathbf{\Sigma}_2 \otimes \mathbf{\Sigma}_1). \tag{5}$$

In this framework, both $\mathcal{N}_1$ and $\mathcal{N}_2$ are regarded as incorrectly-reduced (or mis-specified) models of $\mathcal{M}$ in (1). In order to derive general conclusions, no distributional assumptions are made except the existence of the first and second moments, and no restrictions to the ranks of the matrices $\mathbf{Y}$, $\mathbf{X}$, $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{\Sigma}_1$, and $\mathbf{\Sigma}_2$ in (1)–(5) are required.

A common method of handling MGLMs is to use the well-known Kronecker products and vectorization operators of matrices. Through these operations, (1), (4), and (5) can equivalently be rewritten as the following three UGLMs:

$$\widehat{\mathcal{M}} : \ \overrightarrow{\mathbf{Y}} = (\mathbf{I}_m \otimes \mathbf{X})\overrightarrow{\Theta} + \overrightarrow{\Psi} = (\mathbf{I}_m \otimes \mathbf{X}_1)\overrightarrow{\Theta}_1 + (\mathbf{I}_m \otimes \mathbf{X}_2)\overrightarrow{\Theta}_2 + \overrightarrow{\Psi}, \tag{6}$$

$$\widehat{\mathcal{N}}_1 : \ \overrightarrow{\mathbf{Y}} = (\mathbf{I}_m \otimes \mathbf{X}_1)\overrightarrow{\Theta}_1 + \overrightarrow{\Psi}_1, \tag{7}$$

$$\widehat{\mathcal{N}}_2 : \ \overrightarrow{\mathbf{Y}} = (\mathbf{I}_m \otimes \mathbf{X}_2)\overrightarrow{\Theta}_2 + \overrightarrow{\Psi}_2. \tag{8}$$

This fact demonstrates that there is quite a lot of synergy between UGLMs and MGLMs, and thus we can extend various known facts and results from UGLMs to MGLMs by using the vectorization operations of matrices.

An important problem in the analysis of MGLMs is the search for estimators and the determination of their properties and features under various model assumptions. In this context, statisticians are often interested in describing relationships between different estimators, and especially, in establishing identities between them. Indeed, $\mathbf{X}\Theta = \mathbf{X}_1\Theta_1 + \mathbf{X}_2\Theta_2$ in (1), so that (2)–(5) derive from (1). Hence, inference results on (1)–(8) should be interconnected.

In this paper, we first show that the four fundamental additive decomposition identities

$$\mathrm{OLSE}_{\mathcal{M}}(\mathbf{X}\Theta) = \mathrm{OLSE}_{\mathcal{M}}(\mathbf{X}_1\Theta_1) + \mathrm{OLSE}_{\mathcal{M}}(\mathbf{X}_2\Theta_2),$$
$$\mathrm{BLUE}_{\mathcal{M}}(\mathbf{X}\Theta) = \mathrm{BLUE}_{\mathcal{M}}(\mathbf{X}_1\Theta_1) + \mathrm{BLUE}_{\mathcal{M}}(\mathbf{X}_2\Theta_2),$$
$$\mathrm{OLSE}_{\mathcal{M}}(\mathbf{X}\Theta) = \mathrm{OLSE}_{\mathcal{M}_1}(\mathbf{X}_1\Theta_1) + \mathrm{OLSE}_{\mathcal{M}_2}(\mathbf{X}_2\Theta_2),$$
$$\mathrm{BLUE}_{\mathcal{M}}(\mathbf{X}\Theta) = \mathrm{BLUE}_{\mathcal{M}_1}(\mathbf{X}_1\Theta_1) + \mathrm{BLUE}_{\mathcal{M}_2}(\mathbf{X}_2\Theta_2)$$

always hold under the assumptions that $\mathbf{X}_1\Theta_1$ and $\mathbf{X}_2\Theta_2$ are estimable under $\mathcal{M}$, $\mathcal{M}_1$, and $\mathcal{M}_2$, where the symbols OLSE and BLUE denote the ordinary least squares estimator and the best linear unbiased estimator of unknown parameter matrices under MGLMs, respectively. We then consider the additive decomposition identities

$$\mathrm{OLSE}_{\mathcal{M}}(\mathbf{X}\Theta) = \mathrm{OLSE}_{\mathcal{N}_1}(\mathbf{X}_1\Theta_1) + \mathrm{OLSE}_{\mathcal{N}_2}(\mathbf{X}_2\Theta_2), \tag{9}$$

$$\mathrm{BLUE}_{\mathcal{M}}(\mathbf{X}\Theta) = \mathrm{BLUE}_{\mathcal{N}_1}(\mathbf{X}_1\Theta_1) + \mathrm{BLUE}_{\mathcal{N}_2}(\mathbf{X}_2\Theta_2) \tag{10}$$

under $\mathcal{M}$, $\mathcal{N}_1$, and $\mathcal{N}_2$. These identities have many different statistical interpretations and occur frequently in the statistical analysis of linear regression models. However, (9) and (10) do not necessarily hold in general situations. Thus, we are first interested in establishing necessary and sufficient conditions for the above two additive decomposition equalities to hold.

Additive decompositions of estimators have been an important research topic in the context of classic and recent statistical analysis; see, e.g., [5,7,10,12,18,19,23–25]. As is well known, statistical inference for MGLMs is entirely based on computations with the given matrices in the models, and formulas and algebraic tricks for handling matrices in linear algebra and matrix theory play an important role in the derivations of these estimators and the characterization of their