THE UNIVERSITY OF TEXAS AT SAN ANTONIO, COLLEGE OF BUSINESS

Working Paper SERIES

Date February 5, 2019

WP # 0228MSS

Revisiting the Linear Models with Exchangeably Distributed Errors

Timothy Opheim
The University of Texas at San Antonio
Timothy.Opheim@utsa.edu

Anuradha Roy
The University of Texas at San Antonio
Anuradha.Roy@utsa.edu

Revisiting the Linear Models with Exchangeably Distributed Errors

Timothy Opheim and Anuradha Roy *
Department of Management Science and Statistics
The University of Texas at San Antonio
One UTSA Circle
San Antonio, TX 78249, USA

Abstract

The popularity of the classical general linear model (CGLM) is mostly due to the ease of modeling and authentication of the appropriateness of the model. However, CGLM is not appropriate and thus not applicable for correlated two dimensional observations. In this paper an extension of Arnold's (1979) linear model with exchangeably distributed errors for multiple observations is proposed, the maximum likelihood estimates of the matrix parameters of the intercept, slope and the eigenblocks of the exchangeable error matrix are derived. The practical implications of the methodological aspects of the proposed extended model are demonstrated using two medical datasets.

Keywords block exchangeable covariance structure; linear models; maximum likelihood estimates; hypothesis testing.

JEL Classification C12; C13; C21

1 Introduction

Theoretical inference in statistics is primarily based on the assumption of independent and identically distributed random samples drawn from a population. However, we always do not have samples that are truly random in nature. In these cases the standard statistical inference results fail. Consider an example of a digital image where contiguous pixels are correlated. The correlation exists because sensors take a significant energy from these contiguous pixels, and sensors cover a land region much larger than the size of a pixel. A pixel typically has three or four components (q) such as red, green and blue where each pixel is a point in the RGB space. Thus, RGB space is a 3-dimensional vector space, and each pixel is defined by an ordered triple of red, green and blue coordinates. Correlations exist among the intensities of the ordered triple of red, green and blue coordinates. As mentioned before

*Correspondence to: Anuradha Roy, Department of Management Science and Statistics, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA

correlation also exists among the contiguous pixels. For example, if a pixel represents wheat in an agricultural field, then its neighboring pixels also represent wheat with high probability (Richards et al., 1999). So, there are two types of correlations here, one is within the pixel, and the other one is between the neighboring pixels. This dataset has multivariate repeated measurements, three intensities (q) are repeatedly measured over contiguous pixels (n). A model based on samples of these contiguous pixels must take into account of these two types of correlations, and multivariate equicorrelation could be a reasonable assumption. Arnold (1979) developed linear models for $n \times q$ matrix-variate realization of one multivariate equicorrelated sample with block exchangeable covariance structure (defined later) for the error term. Block exchangeable (BE) covariance structure for matrix-variate data is a generalization of the exchangeable covariance structure for vector-variate data and has been studied most extensively by Arnold (1976) and Szatrowski (1976). Matrix-variate data, where q variables are measured at n locations (sites) or time points are known as multivariate repeated measures data or doubly multivariate data, where the observations in $(n \times q)-$dimensional matrix-variate sample are not independent, but doubly correlated. Uncorrelated error is often a violated assumption of statistical procedures in these kinds of data. Violations occur when error terms are not independent, but instead clustered by one or more grouping variables. Separable covariance structure or BE covariance structure are suitable variance-covariance matrices to analyze matrix-variate observations. While analyzing LANDSAT images Craig (1979) suggested the simplest method of correction for the correlation on the contiguous pixels is by taking samples at increments of ten or more pixels in order to consider them as independent. However, in this way, one fails to take into account the correlation that truly exists in the contiguous pixels. A much better method would be to take some contiguous pixels of samples, and then skip ten or more pixels, and then again take some contiguous pixels of samples. In this way one can get a set of independent contiguous pixels of samples (N), which can preserve both the existing correlations that are present in the data. Another example is from a medical study of cerebral metabolism in epileptic patients (Sperling et al. (1990)). In this study, the metabolic rate of glucose at 16 locations in the brain by positron emission tomography (PET) scans was measured. These 16 locations include 8 regions of interest: the first five, frontal, sensorimotor, temporal, parietal, and occipital are known as cortical regions, and the last three, caudate nucleus, lenticular nucleus and thalamus are known as subcortical regions. Sperling et al. (1990) measured the right-sided (R) and the left-sided (L) metabolic rates in each region of interest. Thus, this study has doubly repeated measurements, namely, observations collected at different locations from the same side of the brain and observations collected in different sides of the brain from the same location. Thus, for these datasets $q = 5$ and $n = 2$, and $q = 3$ and $n = 2$ in cortical and subcortical regions, respectively. The sample consisted of 18 Normal control subjects, 8 patients with a Left brain hemisphere focus of the epilepsy, and 8 patients with a Right brain hemisphere focus of the epilepsy. We use part of this data as an illustration of our method, namely Left hemisphere focus group and Right hemisphere focus group of the cortical region.

As mentioned before, one may fit multivariate linear model for multivariate repeated measures data, or doubly multivariate data using separable error structure $\mathbf{\Psi} \otimes \mathbf{\Omega}$, where $\mathbf{\Psi}$ ($n \times n$ positive definite) and $\mathbf{\Omega}$ ($q \times q$ positive definite) are unstructured (UN) variance-covariance matrices on both the multivariate levels (say, location and response variable). The symbol $\otimes$ represents the Kronecker product. Separable structure sas been studied by many authors (Dutilleul, 1999; Lu and Zimmerman, 2005; Srivastava et al. 2008; Roy and Khattree, 2003). Roy and Khattree (2005a,b) also studied this separable structure by assuming a compound symmetry (CS) or an autoregressive of order one (AR(1)) correlation structure on $\mathbf{\Psi}$.

Mixed procedure of the statistical software package SAS (2009) can fit linear model for doubly multivariate data with separable covariance structures with both unstructured, and half structured and half unstructured variance-covariance matrices as the components of the Kronecker product using a REPEATED statement to identify the correlated observations and to impose structure on the variance-covariance matrix of the response variables. Using SAS procedure *PROC MIXED*, Roy (2006) developed classification rule in conjunction with separable covariance structure, with the first component as CS or AR(1) correlation structure, as the residual variance-covariance matrix for doubly multivariate data. Regrettably, *PROC MIXED* cannot fit linear model for doubly multivariate data with the BE covariance structure.

The aim of the paper is to develop a linear model for doubly multivariate data with the BE covariance structure for multiple (replicated) $n \times q$ matrix-variate observations (N) and develop hypotheses testing procedures to test the intercept vector and the slope matrix. We also derive the maximum likelihood estimates (MLEs) of the model parameters to perform tests of hypotheses for the intercept and the slope parameters. In doing so, we first revisit Arnold's linear model for $n \times q$ matrix-variate realization of one matrix-variate observation (multivariate equicorrelated sample), and generalize his model for N independent multivariate equicorrelated samples, where each independent sample has $(n \times q)-$dimensional matrix-variate observation. To the best of our knowledge, the proposed extended method in this paper is the first study to fit linear model for multiple matrix-variate observations with BE covariance structure. Furthermore, traditional linear models fail for the high-dimensional ($nq > N$) data. Nonetheless, our new generalized EGLM can fit linear models for high-dimensional data and also able to perform hypotheses testing for the model parameters.

The classical linear models for the multivariate data with a vector valued random variable (response vector) can be extended to a multivariate repeated measures data or doubly multivariate data with a matrix valued random variable $\underset{n\times q}{\boldsymbol{Y}}$. For example, classical general linear model (CGLM) for doubly multivariate data can be presented as

$$\underset{n\times q}{\boldsymbol{Y}} = \underset{n\times r}{\boldsymbol{X}}\;\underset{r\times q}{\boldsymbol{B}} + \underset{n\times q}{\boldsymbol{E}},$$

where $\boldsymbol{X}$ is a $n \times r$ design matrix, $\boldsymbol{B}$ is an $r \times q$ matrix of unknown constants and $\boldsymbol{E}$ is the $n \times q$ error matrix.

If the n rows of $\boldsymbol{Y}$ are exchangeable (see e.g., Arnold 1973, 1979 and Žežula, et al. 2018), then $E(\boldsymbol{Y}) = \boldsymbol{j}_n\boldsymbol{\alpha}$ (constant mean vector structure over sites or positions), where $\boldsymbol{j}_n$ is the $n \times 1$ vector of ones.. Thus, the number of free parameters in the mean vector is only q, and the block exchangeable (BE) or block compound symmetry (BCS) covariance structure of vec$(\boldsymbol{E}')$ is then

$$\begin{aligned}\underset{nq\times nq}{\boldsymbol{\Sigma}} &= \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_1 \end{pmatrix} \\ &= \boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2,\end{aligned}$$

where each column of $\boldsymbol{E}'$ has the variance-covariance matrix $\boldsymbol{\Sigma}_1$, and any two different columns have the covariance matrix $\boldsymbol{\Sigma}_2$. Here $\boldsymbol{I}_n$ is the $n \times n$ identity matrix, $\boldsymbol{J}_n = \boldsymbol{j}_n\boldsymbol{j}_n'$. The 'vec' operator is a vectorization of a matrix created by stacking the columns of a matrix on top of one another. In other words, vec of a matrix is a linear transformation which converts the matrix into a column vector. Exchangeability is imposed on $\boldsymbol{Y}$ for considerations that are external to the data – e.g., design considerations of the data. So, the covariance matrix is tied to the same invariance, that is why BE covariance structure is appropriate.

We assume that the $q \times q$ matrix $\boldsymbol{\Sigma}_1$ is positive definite (denoted by $\boldsymbol{\Sigma}_1 > 0$), and the $q \times q$ symmetric matrix $\boldsymbol{\Sigma}_2$ must satisfy $\boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2 > 0$ and $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2 > 0$ (which means that $\boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$ are positive definite matrices) in order to assure the positive definiteness of $\boldsymbol{\Sigma}$ (for a proof, see Lemma 2.1 in Roy and Leiva (2011)).

The pattern of the BE covariance structure arises from imposing symmetry on blocks of variables. From the definition of $\boldsymbol{\Sigma}$, it is clear that $n \geq 2$ is needed for the BE covariance structure, which is a realistic assumption in many doubly multivariate data, and it substantially reduces the number of unknown parameters. A $(nq \times nq)-$dimensional unstructured variance-covariance matrix has $nq(nq+1)/2$ unknown parameters, which can be large for arbitrary values of n or q, whereas the BE structure has only $q(q+1)$ unknown parameters, which is much less than the unstructured one, thus it offers more reliable estimates of the unknown parameters. Additionally, this number does not even depend on n. This means that one can get more information by increasing n, without estimating more parameters, and thus BE covariance structure is appropriate for the high-dimensional matrix-variate data.

The rest of the article is organized as follows. We first set up some preliminaries about exponential class, completeness, sufficiency, and uniformly minimum-variance unbiased estimator (UMVUE) and briefly describe some properties of the matrix-variate normal distribution in Section 2. We then revisit Arnold's model in Section 3. We derive the pseudo maximum likelihood estimates (MLEs) of Arnold's model parameters and discuss their statistical properties in a formal way in Section 4. The extension of Arnold's model for multiple observations is implemented in Section 5 together with the technical derivation of the MLEs of it's matrix parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$. Distribution of the matrix parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for sample of size N are given in Section 6. The joint complete sufficient statistics for the model parameters for multiple observations are derived in Section 7. The tests of hypotheses of intercept and slope parameters are developed in Section 8. Some real data examples are given in Section 9, and finally, Section 10 concludes with some remarks and the scope for the future research.

2 Preliminaries

2.1 Exponential class, completeness, sufficiency, and UMVUE

Suppose $\boldsymbol{x}$ is a $q-$dimensional random vector with probability density function (pdf) $f(\boldsymbol{x};\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \boldsymbol{\Omega} \subset \Re^k$. Let the support of $\boldsymbol{x}$ be a subset of $\Re^q$. Suppose $f(\boldsymbol{x};\boldsymbol{\theta})$ is of the form

$$f(\boldsymbol{x};\boldsymbol{\theta}) = \begin{cases} \exp\Big(\sum_{j=1}^{m} \operatorname{tr}\big(\boldsymbol{p}_j(\boldsymbol{\theta})\boldsymbol{K}_j(\boldsymbol{x})\big) + \boldsymbol{H}(\boldsymbol{x}) + \boldsymbol{q}(\boldsymbol{\theta})\Big) & \forall\, \boldsymbol{x} \in \boldsymbol{S}, \\ \boldsymbol{0} & \text{elsewhere.} \end{cases} \tag{2.1}$$

Suppose $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ constitute a random sample on $\boldsymbol{x}$, then define

$$y_j = \sum_{i=1}^{n} \boldsymbol{K}_j(\boldsymbol{x}_i) \text{ for } j = 1, 2, \ldots, m.$$

We say $f(\boldsymbol{x};\boldsymbol{\theta})$ is a member of the exponential class if it can be expressed in the form defined by (2.1). If, in addition, the dimensions of the $(y_1, y_2, \ldots, y_m)$ equals k, the dimension of the parameter space, and the support of $f(\boldsymbol{x};\boldsymbol{\theta})$ does not depend on the parameters $\boldsymbol{\theta}$, then we say this pdf is a regular case of the exponential class. Moreover, the collection of y_j's is a joint complete sufficient statistic for $\boldsymbol{\theta}$ if $n \geq m$.

Let $\boldsymbol{Y} = (y_1, y_2, \ldots, y_m)'$, and suppose $\delta = g(\boldsymbol{\theta})$ is a parameter of interest. If $T = h(\boldsymbol{Y})$ for some function h and $E(T) = \delta$, then T is the unique minimum variance unbiased estimator (UMVUE) of δ.

3 The model

As mentioned in the Introduction, multivariate repeated measures data, or doubly multivariate data, are data where the observations in each $(n \times q)-$dimensional matrix-variate sample are doubly correlated. For example, analysis of multivariate repeated measures data needs to take into account the correlations among the measurements of q different variables as well as the correlations among measurements taken at n different locations or time points. Summation of observations is not appropriate as it results in a loss of detailed information of particular characteristics that may be of interest. The ability to analyze the multivariate repeated measurement data on its original scale may therefore be advantageous.

Arnold (1979) extended the classical linear models for the multivariate data with a vector-valued random variable (response vector) to linear model with BE error matrix for the multivariate repeated measures data or doubly multivariate data with a matrix-valued random variable $\boldsymbol{Y}$, which he named as exchangeable general linear model (EGLM). We resume the CGLM in the Section 3.1 and revisit the extension of it, the EGLM, in Section 3.2.

3.1 Classical general linear model

Suppose we obtain a sample of size n such that each datum has q response variables and $r-1$ predictor variables associated with it. Let $\boldsymbol{Y}$ be an $(n \times q)$ random matrix which represents the matrix of responses, $\boldsymbol{\alpha}$ be a $(1 \times q)$ vector, $\boldsymbol{T}$ be an $(n \times (r-1))$ matrix of constants, $\boldsymbol{\gamma}$ be an $\big((r-1) \times q\big)$ matrix of parameters, and $\boldsymbol{E}$ be an $(n \times q)$ random matrix which represents the matrix of errors. Also let $\boldsymbol{y}_i'$ be the $i-$th row vector of $\boldsymbol{Y}$, $\boldsymbol{t}_i'$ be the $i-$th row vector of $\boldsymbol{T}$, and $\boldsymbol{e}_i'$ be the $i-$th row vector of $\boldsymbol{E}$. The CGLM for doubly multivariate data $\boldsymbol{Y}$ can be presented as

$$\begin{aligned}
\begin{pmatrix} \underset{1\times q}{\boldsymbol{y}_1'} \\ \vdots \\ \underset{1\times q}{\boldsymbol{y}_n'} \end{pmatrix} &= \begin{bmatrix} \underset{1\times q}{\boldsymbol{\alpha}} + \boldsymbol{t}_1'\gamma \\ \vdots \\ \underset{1\times q}{\boldsymbol{\alpha}} + \boldsymbol{t}_n'\gamma \end{bmatrix} + \begin{pmatrix} \underset{1\times q}{\boldsymbol{e}_1'} \\ \vdots \\ \underset{1\times q}{\boldsymbol{e}_n'} \end{pmatrix}, \\
\underset{n\times q}{\boldsymbol{Y}} &= \boldsymbol{j}_n\boldsymbol{\alpha} + \underset{n\times(r-1)}{\boldsymbol{T}}\gamma + \boldsymbol{E} \qquad (3.2) \\
&= \underset{n\times r}{\boldsymbol{X}}\,\underset{r\times q}{\boldsymbol{B}} + \underset{n\times q}{\boldsymbol{E}},
\end{aligned}$$

where the design matrix $\boldsymbol{X}$ and the matrix parameter $\boldsymbol{B}$ are

$$\boldsymbol{X} = \big[\boldsymbol{j}_n : \boldsymbol{T}\big] \quad \text{and} \quad \underset{r\times q}{\boldsymbol{B}} = \begin{pmatrix} \underset{1\times q}{\boldsymbol{\alpha}} \\ \underset{r-1\times q}{\gamma} \end{pmatrix},$$

with full column rank of $\boldsymbol{X}$. The CGLM makes the following assumptions about the rows of the random error matrix $\boldsymbol{E}$, where

$$\underset{n\times q}{\boldsymbol{E}} = \begin{pmatrix} \boldsymbol{e}_1' \\ \vdots \\ \boldsymbol{e}_n' \end{pmatrix}, \qquad (3.3)$$

with $E(\boldsymbol{e}_i) = \underset{q\times 1}{\boldsymbol{0}}$ and $E(\boldsymbol{e}_i\boldsymbol{e}_i') = \underset{q\times q}{\boldsymbol{\Sigma}_1}$ $\forall\ i = 1, 2, \ldots, n,$ and $E(\boldsymbol{e}_i\boldsymbol{e}_{i^*}') = \underset{q\times q}{\boldsymbol{0}}$ $\forall\ i, i^* = 1, 2, \ldots, n,\ i \neq i^*$ for some unknown $q \times q$ symmetric positive definite matrix $\boldsymbol{\Sigma}_1$.

Now,

$$E\big[\text{vec}(\boldsymbol{E}')\big] = E\begin{pmatrix}\boldsymbol{e}_1\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix} = \begin{pmatrix}\mathbf{0}_q\\ \vdots\\ \mathbf{0}_q\end{pmatrix} = \text{vec}(\underset{q\times n}{\mathbf{0}}) \quad \text{and} \tag{3.4}$$

$$\text{Cov}\big[\text{vec}(\boldsymbol{E}')\big] = E\left[\begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix}(\boldsymbol{e}_1', \boldsymbol{e}_1', \dots, \boldsymbol{e}_n')\right] = \begin{pmatrix}\boldsymbol{\Sigma}_1 & \underset{q\times q}{\mathbf{0}} & \dots & \underset{q\times q}{\mathbf{0}}\\ \underset{q\times q}{\mathbf{0}} & \boldsymbol{\Sigma}_1 & \dots & \underset{q\times q}{\mathbf{0}}\\ \vdots & \vdots & \ddots & \vdots\\ \underset{q\times q}{\mathbf{0}} & \underset{q\times q}{\mathbf{0}} & \dots & \boldsymbol{\Sigma}_1\end{pmatrix} = \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_1.$$

Since the collection of all $\boldsymbol{e}_i'$ are independent and jointly normally distributed, we have

$$\boldsymbol{E} \sim N_{n,q}(\underset{n\times q}{\mathbf{0}}, \boldsymbol{I}_n, \boldsymbol{\Sigma}_1), \text{ or}$$
$$\text{vec}(\boldsymbol{E}') \sim N_{nq}(\text{vec}(\underset{q\times n}{\mathbf{0}}), \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_1).$$

Therefore, the columns of $\boldsymbol{E}'$ are independent and identically distributed

$$\boldsymbol{e}_i \sim N_q(\mathbf{0}_q, \boldsymbol{\Sigma}_1), \forall\ i = 1, 2, \dots, n.$$

Hence the columns of $\boldsymbol{Y}'$, $\boldsymbol{y}_i$, are independent and $q-$variate normally distributed as follows

$$\boldsymbol{y}_i \sim N_q(\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{t}_i, \boldsymbol{\Sigma}_1), \forall\ i = 1, 2, \dots, n.$$

3.2 Exchangeable general linear model

The EGLM is still defined by the model (3.2), nonetheless it makes the following assumptions about the rows of the random error matrix $\boldsymbol{E}$ defined in (3.3):
$E(\boldsymbol{e}_i) = \underset{q\times 1}{\mathbf{0}}$ and $E(\boldsymbol{e}_i\boldsymbol{e}_i') = \underset{q\times q}{\boldsymbol{\Sigma}_1}\ \forall\ i = 1, 2, \dots, n,$ and $E(\boldsymbol{e}_i\boldsymbol{e}_{i^*}') = \underset{q\times q}{\boldsymbol{\Sigma}_2}\ \forall\ i, i^* = 1, 2, \dots, n,\ i \neq i^*$ for some unknown $q \times q$ symmetric positive definite matrix $\boldsymbol{\Sigma}_1$ and for some unknown $q \times q$ symmetric matrix $\boldsymbol{\Sigma}_2$, so that the joint covariance matrix of $\boldsymbol{e}_i$ for $i = 1, 2, \dots, n$ is positive definite. See Lemma 2.1 in Roy and Leiva (2011). Collectively, these assumptions imply the expectation matrix of vec$(\boldsymbol{E}')$ be the same as in Equation (3.4), and the variance-covariance matrix of vec$(\boldsymbol{E}')$ will be as follows:

$$\text{Cov}\big[\text{vec}(\boldsymbol{E}')\big] = E\left[\begin{pmatrix}\boldsymbol{e}_1\\ \boldsymbol{e}_2\\ \vdots\\ \boldsymbol{e}_n\end{pmatrix}(\boldsymbol{e}_1', \boldsymbol{e}_1', \dots, \boldsymbol{e}_n')\right] = \begin{pmatrix}\boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_2\\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_2\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_1\end{pmatrix}$$
$$= \boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2. \tag{3.5}$$

Since the collection of all $\boldsymbol{e}_i'$ are exchangeable and jointly normally distributed, using the properties of matrix-variate normal distribution we have

$$\begin{aligned} \boldsymbol{E}' &\sim N_{q,n}(\underset{q\times n}{\mathbf{0}}, \boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2), \text{ or} \\ \text{vec}(\boldsymbol{E}') &\sim N_{nq}(\text{vec}(\underset{q\times n}{\mathbf{0}}), \boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2). \end{aligned}$$

3.2.1 Eigendecomposition of block exchangeable covariance structure

Let $\boldsymbol{C}^* = \boldsymbol{C}' \otimes \boldsymbol{I}_q$, where $\boldsymbol{C}$ is an $n \times n$ orthogonal Helmert matrix whose first column is proportional to a vector of 1's. So, $\boldsymbol{C}^*$ is an orthogonal matrix too. Now pre-multiplying the $\text{Cov}\big[\text{vec}(\boldsymbol{E}')\big]$ in (3.5) by $\boldsymbol{C}^*$ and post-multiplying it by its transpose we get the following block diagonal matrix:

$$\boldsymbol{C}^*[\boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2]\boldsymbol{C}^{*\prime} = \begin{bmatrix} \boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{n-1} \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) \end{bmatrix}. \tag{3.6}$$

The positive definite matrices $\boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$ are two distinct eigenblocks (see Hao et al. 2015, Arnold 1973) of the BE covariance structure (3.5) with $(n-1)$ repetitions of the second eigenblock $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$.

3.2.2 The transformed model

Arnold (1979) used the eigendecomposition of BE covariance structure in his favor and rotated the EGLM by pre-multiplying it by the orthogonal matrix $\boldsymbol{C}'$. In this way he transformed the EGLM into n independent CGLMs (basically the principal vectors), where the variance-covariance matrices of the error terms of these n CGLMs are the n eigenblocks of the BE structure (3.5). The transformed model is given by

$$\boldsymbol{C}'\boldsymbol{Y} = \boldsymbol{C}'\boldsymbol{j}_n\boldsymbol{\alpha} + \boldsymbol{C}'\boldsymbol{T}\boldsymbol{\gamma} + \boldsymbol{C}'\boldsymbol{E}.$$

Hence,

$$\begin{pmatrix} \boldsymbol{z}_1 \\ \boldsymbol{Z}_2 \end{pmatrix} = \begin{pmatrix} \sqrt{n}\boldsymbol{\alpha} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \boldsymbol{u}_1\boldsymbol{\gamma} \\ \boldsymbol{U}_2\boldsymbol{\gamma} \end{pmatrix} + \begin{pmatrix} \boldsymbol{E}_1 \\ \boldsymbol{E}_2 \end{pmatrix},$$

where

$$\boldsymbol{C}'\boldsymbol{Y} = \begin{pmatrix} \underset{1\times q}{\boldsymbol{z}_1} \\ \underset{(n-1)\times q}{\boldsymbol{Z}_2} \end{pmatrix}, \quad \boldsymbol{C}'\boldsymbol{T} = \begin{pmatrix} \underset{1\times r-1}{\boldsymbol{u}_1} \\ \underset{n-1\times r-1}{\boldsymbol{U}_2} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{C}'\boldsymbol{E} = \begin{pmatrix} \underset{1\times q}{\boldsymbol{E}_1} \\ \underset{(n-1)\times q}{\boldsymbol{E}_2} \end{pmatrix}.$$

Denote the $j-$th row vector of $\boldsymbol{U}_2$ as $\boldsymbol{u}'_{2,j}$ and the $j-$th row vector of $\boldsymbol{Z}_2$ as $\boldsymbol{z}'_{2,i}$

$$\underset{n-1\times q}{\boldsymbol{Z}_2} = \begin{pmatrix} \boldsymbol{z}'_{2,1} \\ \vdots \\ \boldsymbol{z}'_{2,n-1} \end{pmatrix}, \quad \text{and} \quad \underset{n-1\times r-1}{\boldsymbol{U}_2} = \begin{pmatrix} \boldsymbol{u}'_{2,1} \\ \vdots \\ \boldsymbol{u}'_{2,n-1} \end{pmatrix}. \tag{3.7}$$

Using the eigendecomposition results in Section 3.2.1 and the basic properties of multivariate normal distribution, we see $\boldsymbol{z}'_1$ and $\text{vec}(\boldsymbol{Z}'_2)$ are independent with distributions as follows

$$\boldsymbol{z}'_1 \sim N_q\big(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}'_1, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\big) \quad \text{and} \tag{3.8}$$

$$\text{vec}(\boldsymbol{Z}'_2) \sim N_{(n-1)q}\big(\text{vec}(\boldsymbol{\gamma}'\boldsymbol{U}'_2), \boldsymbol{I}_{n-1} \otimes \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big), \tag{3.9}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{z}_1} = \boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2} = \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$. Hence,

$$\boldsymbol{z}_{2,j} \sim N_q(\boldsymbol{\gamma}'\boldsymbol{u}_{2,j}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}), \quad \text{for } j = 1, 2, \ldots, n-1.$$

According to Arnold, "The importance of the transformed model is that the EGLM has been transformed into two separate models." In the model involving $\boldsymbol{z}_1$, there is no replication; so estimation of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is not possible. No replication means estimates of the variances of all q transformed variables are zeros and estimates of the covariances among them are zeros too. This suggests the estimate of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is a matrix with all zero entries. Nevertheless, the model involving $\boldsymbol{Z}_2$ has $(n-1)$ replications with no intercept term. This model is just a CGLM, so it is possible to estimate $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$.

Arnold (1979) obtained the least square estimates of the model parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$, and he mentioned that there were no MLEs for the EGLM, but only pseudo-MLEs for fixed $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$, and no sensible estimators exists for $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. We derive the pseudo-MLEs of the EGLM parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ for fixed $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ in the following section.

4 Pseudo-maximum likelihood estimation of transformed EGLM for one observation

Theorem 1 *The pseudo-MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ in the EGLM (3.2) for fixed $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ are given by*

$$\widehat{\boldsymbol{\alpha}}' = \frac{\boldsymbol{z}'_1 - \widehat{\gamma}'\boldsymbol{u}'_1}{\sqrt{n}}, \quad \textit{and}$$

$$\widehat{\gamma} = (\boldsymbol{U}'_2\boldsymbol{U}_2)^{-1}\boldsymbol{U}'_2\boldsymbol{Z}_2.$$

and these pseudo-MLEs are exactly the same as the least square estimates of $\boldsymbol{\alpha}'$ *and* $\boldsymbol{\gamma}$ *respectively obtained by Arnold (1979).*

Proof 1 *In order to avoid the matrix normal density, all estimation will be carried out on maximizing the likelihood equations of* $\boldsymbol{z}_1'$ *and* $vec(\boldsymbol{Z}_2')$ *so as to make use of known results from the multivariate normal density. Moreover, since* $\boldsymbol{\alpha}'$ *only appears in* $\boldsymbol{z}_1'$*, we can ignore the density of* $vec(\boldsymbol{Z}_2')$ *in the estimation of* $\boldsymbol{\alpha}'$*.*

Since $\sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1'$ *and* $\boldsymbol{z}_1'$ *are* $q \times 1$ *vector,*

$$vec(\sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1') \quad = \quad \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1'.$$

Therefore, the likelihood equation of $\boldsymbol{z}_1'$ *given* $\boldsymbol{\gamma}$ *and* $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ *have been estimated as*

$$L(\boldsymbol{\alpha}'|\boldsymbol{z}_1', \widehat{\gamma}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}) = (2\pi)^{-\frac{q}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \widehat{\gamma}'\boldsymbol{u}_1')'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \widehat{\gamma}'\boldsymbol{u}_1')\right\},$$

The kernel of the likelihood equation involving $\boldsymbol{\alpha}'$ *is*

$$L(\boldsymbol{\alpha}'|\boldsymbol{z}_1', \widehat{\gamma}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}) = exp\left\{\sqrt{n}\boldsymbol{z}_1\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \frac{1}{2}n\boldsymbol{\alpha}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \sqrt{n}\boldsymbol{u}_1\widehat{\gamma}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}'\right\}.$$

Hence the log of the kernel of this likelihood equation is

$$\ln L(\boldsymbol{\alpha}'|\boldsymbol{z}_1', \widehat{\gamma}, \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}) = \sqrt{n}\boldsymbol{z}_1\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \frac{1}{2}n\boldsymbol{\alpha}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \sqrt{n}\boldsymbol{u}_1\widehat{\gamma}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}'.$$

Taking the partial derivative of the above log likelihood kernel with respect to (w.r.t.) $\boldsymbol{\alpha}'$ *and equating it to* $\mathbf{0}_q$ *we get*

$$\sqrt{n}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1' - n\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \sqrt{n}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}\widehat{\gamma}'\boldsymbol{u}_1' = \mathbf{0}_q.$$

After some algebraic calculations we get

$$\widehat{\boldsymbol{\alpha}}' = \frac{\boldsymbol{z}_1' - \widehat{\gamma}'\boldsymbol{u}_1'}{\sqrt{n}}.$$

Next we will derive the pseudo-MLE of $\boldsymbol{\gamma}$*. Using the definition of* $\boldsymbol{Z}_2$ *and* $\boldsymbol{U}_2$ *in (3.7) we have*

$$vec(\boldsymbol{\gamma}'\boldsymbol{U}_2') = vec(\boldsymbol{\gamma}'\boldsymbol{u}_{2,1}', \ldots, \boldsymbol{\gamma}'\boldsymbol{u}_{2,n-1}') = \begin{pmatrix} \boldsymbol{\gamma}'\boldsymbol{u}_{2,1} \\ \vdots \\ \boldsymbol{\gamma}'\boldsymbol{u}_{2,n-1} \end{pmatrix}.$$

From the estimation of $\boldsymbol{\alpha}'$*, note that*

$$\boldsymbol{z}_1' - \sqrt{n}\widehat{\boldsymbol{\alpha}}' + \boldsymbol{\gamma}'\boldsymbol{u}_1' = \boldsymbol{0}_q.$$

Hence, maximizing the full likelihood equation w.r.t. $\boldsymbol{\gamma}$ *is equivalent to maximizing the likelihood equation of* $vec(\boldsymbol{Z}_2')$ *alone, since* $\boldsymbol{\gamma}$ *no longer appears in the likelihood equation of* $\boldsymbol{z}_1'$ *after estimating* $\boldsymbol{\alpha}'$*. The likelihood equation of* $vec(\boldsymbol{Z}_2')$ *is*

$$L(\boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|vec(\boldsymbol{Z}_2')) = \prod_{j=1}^{n-1}(2\pi)^{-\frac{q}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{z}_{2,j} - \boldsymbol{\gamma}'\boldsymbol{u}_{2,j})'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{z}_{2,j} - \boldsymbol{\gamma}'\boldsymbol{u}_{2,j})\right\},$$

The kernel of the log-likelihood equation involving $\boldsymbol{\gamma}$ *is*

$$\begin{aligned}\ln L(\boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|vec(\boldsymbol{Z}_2')) &= \sum_{j=1}^{n-1}\boldsymbol{u}_{2,j}'\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{z}_{2,j} - \frac{1}{2}\sum_{j=1}^{n-1}\boldsymbol{u}_{2,j}'\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j} \\ &= \sum_{j=1}^{n-1} tr(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{z}_{2,j}\boldsymbol{u}_{2,j}') - \frac{1}{2}\sum_{j=1}^{n-1} tr(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j}\boldsymbol{u}_{2,j}').\end{aligned} \tag{4.10}$$

Note that

$$\begin{aligned}\boldsymbol{U}_2'\boldsymbol{Z}_2 &= (\boldsymbol{u}_{2,1}, \ldots, \boldsymbol{z}_{2,n-1})\begin{pmatrix}\boldsymbol{z}_{2,1}' \\ \vdots \\ \boldsymbol{z}_{2,n-1}'\end{pmatrix} = \sum_{j=1}^{n-1}\boldsymbol{u}_{2,j}\boldsymbol{z}_{2,j}', \quad \textit{and} \\ \boldsymbol{U}_2'\boldsymbol{U}_2 &= (\boldsymbol{u}_{2,1}, \ldots, \boldsymbol{z}_{2,n-1})\begin{pmatrix}\boldsymbol{u}_{2,1}' \\ \vdots \\ \boldsymbol{u}_{2,n-1}'\end{pmatrix} = \sum_{j=1}^{n-1}\boldsymbol{u}_{2,j}\boldsymbol{u}_{2,j}'.\end{aligned}$$

Using the above results and taking the partial derivative of the above log likelihood kernel (4.10) w.r.t. $\boldsymbol{\gamma}$ *(non-symmetric) and equating it to* $\underset{(r-1)\times q}{\boldsymbol{0}}$ *we get*

$$\begin{aligned}\boldsymbol{U}_2'\boldsymbol{Z}_2\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} - \boldsymbol{U}_2'\boldsymbol{U}_2\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} &= \underset{(r-1)\times q}{\boldsymbol{0}} \quad \textit{or} \\ \boldsymbol{U}_2'\boldsymbol{U}_2\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} &= \boldsymbol{U}_2'\boldsymbol{Z}_2\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} \quad \textit{or} \\ \widehat{\boldsymbol{\gamma}} &= (\boldsymbol{U}_2'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2'\boldsymbol{Z}_2.\end{aligned}$$

Theorem 2 *The pseudo-MLEs of* $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ *in (3.9) for fixed* $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ *is given by*

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} = \frac{(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})}{n-1}, \tag{4.11}$$

however, the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ *in (3.8) does not exist.*

Proof 2 *Since $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2} = \boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$ only appears in vec$(\boldsymbol{Z}_2')$, we can ignore the density of $\boldsymbol{z}_1'$ in the estimation of $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2$. The likelihood equation of vec$(\boldsymbol{Z}_2')$ given $\boldsymbol{\gamma}$ has been estimated as*

$$L(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|vec(\boldsymbol{Z}_2'), \widehat{\boldsymbol{\gamma}}) = \prod_{j=1}^{n-1} (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}} exp\left\{ -\frac{1}{2}(\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j})' \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} (\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j}) \right\},$$

The kernel of the log-likelihood w.r.t $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ is

$$\begin{aligned} \ln L(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|vec(\boldsymbol{Z}_2'), \widehat{\boldsymbol{\gamma}}) &= \frac{n-1}{2} \ln |\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}| - \frac{1}{2} \sum_{j=1}^{n-1} (\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j})' \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} (\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j}) \\ &= \frac{n-1}{2} \ln |\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}| - \frac{1}{2} \sum_{j=1}^{n-1} tr\left(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} (\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j})(\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j})' \right). \end{aligned}$$

Taking the partial derivative of the kernel of the log-likelihood w.r.t. $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}$ (symmetric) and equating it to $\underset{q \times q}{\mathbf{0}}$ and simplifying we get

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} = \frac{1}{n-1} \sum_{j=1}^{n-1} (\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j})(\boldsymbol{z}_{2,j} - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j})'.$$

Expanding the expression within the summation and using the results developed for estimation of $\boldsymbol{\gamma}$, the pseudo-MLE of $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ can be expressed in the matrix notation as

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} &= \frac{1}{n-1} \sum_{j=1}^{n-1} (\boldsymbol{z}_{2,j}\boldsymbol{z}_{2,j}' - 2\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j}\boldsymbol{z}_{2,j}' + \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2,j}\boldsymbol{u}_{2,j}'\widehat{\boldsymbol{\gamma}}) \\ &= \frac{(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})}{n-1}. \end{aligned}$$

An unbiased estimate of $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ is given by

$$\boldsymbol{S} = \frac{n-1}{n-r} \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} = \frac{(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})}{n-r}.$$

This is the same estimate as obtained by Arnold (1979). We will now show analytically that MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ does not exist. Since $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ only appears in $\boldsymbol{z}_1'$, we can ignore the density of vec$(\boldsymbol{Z}_2')$ in the estimation of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$. Recall that after estimating $\boldsymbol{\alpha}$ the exponent involving $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ equals 0; hence the likelihood equation reduces to

$$\begin{aligned} L(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\boldsymbol{z}_1', \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\alpha}}) &= (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(\boldsymbol{z}_1' - \sqrt{n}\widehat{\boldsymbol{\alpha}}' + \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_1')' \boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1} (\boldsymbol{z}_1' - \sqrt{n}\widehat{\boldsymbol{\alpha}}' + \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_1')) \right\} \\ &= (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}}. \end{aligned}$$

Thus, the kernel of the log-likelihood w.r.t. $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is given by

$$\ln L(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\boldsymbol{z}_1', \widehat{\boldsymbol{\gamma}}, \widehat{\boldsymbol{\alpha}}) = -\frac{1}{2}\ln\left(|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\right).$$

So, the likelihood can be made arbitrarily large as $|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}| \to 0$. Therefore, no MLE exist for $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$.

We will now formally discuss the statistical properties of the pseudo-MLEs of the model parameters in the following section.

4.1 Completeness and sufficiency of $\widehat{\boldsymbol{\alpha}}'$, $\widehat{\boldsymbol{\gamma}}$ and $\boldsymbol{S}$

We wish to show that $(\widehat{\boldsymbol{\alpha}}', \widehat{\boldsymbol{\gamma}}, \boldsymbol{S})$ are joint complete sufficient statistics for the model parameters $(\boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$, when $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is known. If $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is not known, then complications in the following technique arise which will be discussed at the the end of this section.

Since the full density can be factored into two independent densities, the joint complete sufficient statistics for $(\boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$ is found by combining the joint complete sufficient statistics for the individual densities. In order to find the joint complete sufficient statistics, one technique is to show that they are members of the regular case of the multivariate exponential class.

Consider the density function of $\boldsymbol{z}_1'$:

$$f_{\boldsymbol{z}_1'}(\boldsymbol{z}_1'|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}) = (2\pi)^{-\frac{q}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1')'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1')\right\},$$

where $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and $\boldsymbol{\gamma}$ are known, the parameter space is a subset of $\Re^q$, Also, the support of $\boldsymbol{z}_1'$ is $\Re^q$. Rearranging the terms in the above density, we see

$$\begin{aligned} f_{\boldsymbol{z}_1'}(\boldsymbol{z}_1'|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}) &= \exp\left\{-\frac{1}{2}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1')'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1') - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}\right|\right\} \\ = \exp&\left\{\mathrm{tr}(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1')\boldsymbol{z}_1) - \frac{1}{2}\boldsymbol{z}_1\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1' - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}\right| - \frac{1}{2}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{u}_1\boldsymbol{\gamma})\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1')\right\} \\ = \quad &\exp\left\{\mathrm{tr}\left(\boldsymbol{p}_1(\boldsymbol{\alpha}')\boldsymbol{K}_1(\boldsymbol{z}_1')\right) + \boldsymbol{H}(\boldsymbol{z}_1') + \boldsymbol{q}(\boldsymbol{\alpha}')\right\}, \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{p}_1(\boldsymbol{\alpha}') &= \boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\left(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1'\right) \\ \boldsymbol{K}_1(\boldsymbol{z}_1') &= \boldsymbol{z}_1 \\ \boldsymbol{H}(\boldsymbol{z}_1') &= \frac{1}{2}\boldsymbol{z}_1\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1' - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}\right| \quad \text{and} \\ \boldsymbol{q}(\boldsymbol{\alpha}') &= \frac{1}{2}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{u}_1\boldsymbol{\gamma})\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1'). \end{aligned}$$

As there is only one observation, $\boldsymbol{Y}_1 = \boldsymbol{z}_1$. Since $\boldsymbol{Y}_1$ is of the same dimension as the parameter space, and the support of $\boldsymbol{z}_1'$ is $\Re^q$ which does not depend on $\boldsymbol{\alpha}'$, we say the pdf of $\boldsymbol{z}_1'$ is a regular case of the exponential class. Moreover, the statistic $\boldsymbol{z}_1$ is a complete sufficient statistic for $\boldsymbol{\alpha}'$ given $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and $\boldsymbol{\gamma}$ are known. Now, consider the density of one independent component of vec$(\boldsymbol{Z}_2'), \boldsymbol{z}_{2,j}$.

$$f_{\boldsymbol{z}_{2,j}}(\boldsymbol{z}_{2,j}|\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}) = (2\pi)^{-\frac{q}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{z}_{2,j} - \boldsymbol{\gamma}'\boldsymbol{u}_{2,j})'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{z}_{2,j} - \boldsymbol{\gamma}'\boldsymbol{u}_{2,j})\right\},$$

where $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ and $\boldsymbol{\gamma}$ are both unknown, the parameter space is a subset of $\Re^{q(q+2r-1)/2}$. Note that $(r-1)q + q(q+1)/2 = q(q+2r-1)/2$. Also, the support of $\boldsymbol{z}_{2,i}$ is $\Re^q$. Rearranging terms in the above density, we see

$$\begin{aligned}
f_{\boldsymbol{z}_{2,j}}(\boldsymbol{z}_{2,j}|\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}) &= \exp\left\{-\frac{1}{2}\boldsymbol{z}_{2,j}'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{z}_{2,j} + \boldsymbol{z}_{2,j}'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j} - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right| - \frac{1}{2}\boldsymbol{u}_{2,j}'\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j}\right\} \\
&= \exp\left\{\mathrm{tr}(-\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{z}_{2,j}\boldsymbol{z}_{2,j}') + \mathrm{tr}(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j}\boldsymbol{z}_{2,j}') - \frac{1}{2}\boldsymbol{z}_1\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{z}_1' - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right| - \frac{1}{2}\boldsymbol{u}_{2,j}'\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j}\right\} \\
&= \exp\Big\{\mathrm{tr}\big(\boldsymbol{p}_1(\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})\boldsymbol{K}_1(\boldsymbol{z}_{2,j}')\big) + \mathrm{tr}\big(\boldsymbol{p}_2(\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})\boldsymbol{K}_2(\boldsymbol{z}_{2,j}')\big) + \boldsymbol{H}(\boldsymbol{z}_{2,j}') + \boldsymbol{q}(\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})\Big\},
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{p}_1(\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}) &= -\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} \\
\boldsymbol{K}_1(\boldsymbol{z}_{2,j}) &= \boldsymbol{z}_{2,j}\boldsymbol{z}_{2,j}' \\
\boldsymbol{p}_2(\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}) &= -\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}' \\
\boldsymbol{K}_2(\boldsymbol{z}_{2,j}) &= \boldsymbol{u}_{2,j}\boldsymbol{z}_{2,j}' \\
\boldsymbol{H}(\boldsymbol{z}_{2,j}) &= -\frac{p}{2}\ln(2\pi) \quad \text{and} \\
\boldsymbol{q}(\boldsymbol{\gamma}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}) &= -\frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right| - \frac{1}{2}\boldsymbol{u}_{2,j}'\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2,j}.
\end{aligned}$$

As there are $(n-1)$ observations in the sample, $\boldsymbol{Y}_1 = \sum_{j=1}^{n-1}\boldsymbol{z}_{2,j}\boldsymbol{z}_{2,j}'$ and $\boldsymbol{Y}_2 = \sum_{j=1}^{n-1}\boldsymbol{u}_{2,j}\boldsymbol{z}_{2,j}'$. Since $\boldsymbol{Y}_1$ is a symmetric $q \times q$ matrix, it has $q(q+1)/2$ nonredundant statistics. Moreover, since $\boldsymbol{Y}_2$ is a $(r-1) \times q$ matrix, it has $(r-1)q$ statistics. Altogether, since $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are functionally independent, they have $q(q+2r-1)/2$ nonredundant statistics which is of the same dimension as the parameter space. Therefore, $\boldsymbol{z}_{2,i}$ is a regular case of the exponential class. Moreover, $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are joint complete sufficient statistics of $(\boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$. Finally, by the introductory remarks, $(\boldsymbol{z}_1, \boldsymbol{Z}_2'\boldsymbol{Z}_2, \boldsymbol{U}_2'\boldsymbol{Z}_2)$ are joint complete sufficient statistics for $(\boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$, when $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is known.

When $\mathbf{\Sigma}_{\boldsymbol{z}_1}$ is not known, the density of $\boldsymbol{z}_1$ can still be expressed in the form of (2.1); however, it is no longer a regular case of the exponential class nor a curved case of the exponential class, but an over-parameterized case of the exponential class, and we have

$$f_{\boldsymbol{z}_1'}(\boldsymbol{z}_1'|\boldsymbol{\alpha},\boldsymbol{\gamma},\mathbf{\Sigma}_{\boldsymbol{z}_1}) = \exp\left\{-\frac{1}{2}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1')'\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}_1' - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}_1') - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\mathbf{\Sigma}_{\boldsymbol{z}_1}\right|\right\}$$

$$=\exp\left\{\mathrm{tr}(\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1')\boldsymbol{z}_1) + \mathrm{tr}\left(\frac{1}{2}\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1'\boldsymbol{z}_1\right) - \frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\mathbf{\Sigma}_{\boldsymbol{z}_1}\right| - \frac{1}{2}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{u}_1\boldsymbol{\gamma})\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1')\right\}$$

$$= \exp\left\{\mathrm{tr}\left(\boldsymbol{p}_1(\boldsymbol{\alpha}',\mathbf{\Sigma}_{\boldsymbol{z}_1}) + \mathrm{tr}(\boldsymbol{p}_2(\boldsymbol{\alpha}',\mathbf{\Sigma}_{\boldsymbol{Z}_1})\boldsymbol{K}_2(\boldsymbol{z}_1'))\right) + \boldsymbol{H}(\boldsymbol{z}_1') + \boldsymbol{q}(\boldsymbol{\alpha}',\mathbf{\Sigma}_{\boldsymbol{z}_1})\right\},$$

where

$$\begin{aligned}
\boldsymbol{p}_1(\boldsymbol{\alpha}',\mathbf{\Sigma}_{\boldsymbol{z}_1}) &= -\frac{1}{2}\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1') \\
\boldsymbol{K}_1(\boldsymbol{z}_1) &= \boldsymbol{z}_1 \\
\boldsymbol{p}_2(\boldsymbol{\alpha}',\mathbf{\Sigma}_{\boldsymbol{z}_1}) &= -\frac{1}{2}\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1} \\
\boldsymbol{K}_2(\boldsymbol{z}_1) &= \boldsymbol{z}_1'\boldsymbol{z}_1 \\
\boldsymbol{H}(\boldsymbol{z}_1') &= -\frac{q}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{\Sigma}_{\boldsymbol{z}_1}| \quad \text{and} \\
\boldsymbol{q}(\gamma',\mathbf{\Sigma}_{\boldsymbol{z}_1}) &= -\frac{1}{2}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{u}_1\boldsymbol{\gamma})\mathbf{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_1').
\end{aligned}$$

As there is only one observation, $\boldsymbol{Y}_1 = \boldsymbol{z}_1$ and $\boldsymbol{Y}_2 = \boldsymbol{z}_1'\boldsymbol{z}_1$. Since the parameter space has $q + q(q+1)/2$ dimensions, but $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are functionally dependent, they only define a $q-$dimensional space. Since these dimensions do not agree, the properties of joint complete sufficient statistics of the exponential class do not apply. However, using the definition of sufficiency, when one has a single observation, that observation is a joint sufficient statistic for the parameter space. Therefore a joint sufficient statistic for $(\boldsymbol{\alpha}',\mathbf{\Sigma}_{\boldsymbol{z}_1})$ is $\boldsymbol{Y}_1$. Moreover, from the definition of completeness, $\boldsymbol{Y}_1$ is complete since it is a $q-$variate normally distributed with unknown mean vector and variance-covariance matrix. Therefore, we have the following remark.

Remark 1 *Although the preceding remarks are sparse, they may be more illuminating than Arnold's comment concerning sufficiency and completeness. "It is known that the distribution of* $(\widehat{\boldsymbol{\alpha}}',\widehat{\boldsymbol{\gamma}},\boldsymbol{S})$ *is complete for the subfamily of distributions in which* $\mathbf{\Sigma}_2 = \underset{q\times q}{\mathbf{0}}$ *(i.e., the CGLM). Since the distributions are mutually absolutely continuous, the distribution of* $(\widehat{\boldsymbol{\alpha}}',\widehat{\boldsymbol{\gamma}},\boldsymbol{S})$ *is also complete for the larger family of distributions for the EGLM."*

We now extend Arnold's EGLM for N multiple independent observations in the following section.

5 Extension of Arnold's EGLM for multiple observations

We will first define the model for N multiple multivariate equicorrelated $n \times q$ matrix-variate samples observations. We will also derive the MLEs of all the model parameters. We define

$$\underset{n\times r-1}{\boldsymbol{T}_p} = \begin{pmatrix} \boldsymbol{t}'_{1,p} \\ \vdots \\ \boldsymbol{t}'_{n,p} \end{pmatrix}, \quad \text{and} \quad \underset{n\times q}{\boldsymbol{Y}_p} = \begin{pmatrix} \boldsymbol{y}'_{1,p} \\ \vdots \\ \boldsymbol{y}'_{n,p} \end{pmatrix}, \quad \text{for} \quad p = 1, 2, \ldots, N,$$

where $\boldsymbol{t}_{i,p}$ and $\boldsymbol{y}_{i,p}$ respectively are $r \times 1$ and $q \times 1$ dimensional vectors for $i = 1, 2, \ldots, n$. Let

$$\begin{aligned} \underset{nq\times 1}{\boldsymbol{y}^*_p} &= \text{vec}(\boldsymbol{Y}'_p) = \begin{pmatrix} \boldsymbol{y}_{1,p} \\ \vdots \\ \boldsymbol{y}_{n,p} \end{pmatrix}, \quad \text{for} \quad p = 1, 2, \ldots, N, \quad \text{and} \\ \underset{nq\times N}{\boldsymbol{Y}} &= (\boldsymbol{y}^*_1, \boldsymbol{y}^*_2, \ldots, \boldsymbol{y}^*_N). \end{aligned}$$

The normality of the data means

$$\text{vec}\,\boldsymbol{Y} \sim N_{Nnq}\left(\begin{pmatrix} \text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_1) \\ \vdots \\ \text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_N) \end{pmatrix}, \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}\right).$$

By multivariate normality properties, $\boldsymbol{y}^*_p \sim N_{nq}\big(\text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_p), \boldsymbol{\Sigma}\big)$ for $p = 1, 2, \ldots, N$ and $\boldsymbol{y}^*_p$ is independent of $\boldsymbol{y}^*_{p^*}$ for $p, p^* = 1, 2, \ldots, N$ and $p \neq p^*$.

So, the model for N multiple multivariate equicorrelated samples becomes

$$\underset{nq\times N}{\boldsymbol{Y}} = \big(\text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_1) \ldots \text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_N)\big) + \big(\text{vec}(\boldsymbol{E}'_1) \ldots \text{vec}(\boldsymbol{E}'_N)\big), \tag{5.12}$$

where

$$\text{vec}\,\boldsymbol{Y} \sim N_{nq,N}\left(\big(\text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_1) \ldots \text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}'_n + \boldsymbol{\gamma}'\boldsymbol{T}'_N)\big), \boldsymbol{\Sigma}, \boldsymbol{I}_N\right).$$

Let the $n \times n$ orthogonal Helmert matrix $\boldsymbol{C} = (\frac{\boldsymbol{j}_n}{\sqrt{n}}, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_n)$, and as defined before $\boldsymbol{C}^* = \boldsymbol{C}' \otimes \boldsymbol{I}_q$. Then

$$\boldsymbol{C}'\boldsymbol{Y}_p = \begin{pmatrix} \boldsymbol{z}_{1,p} \\ \boldsymbol{Z}_{2,p} \end{pmatrix} = \begin{pmatrix} \frac{\boldsymbol{j}'_n}{\sqrt{n}}\boldsymbol{Y}_p \\ \begin{pmatrix} \boldsymbol{c}'_2 \\ \vdots \\ \boldsymbol{c}'_n \end{pmatrix} \boldsymbol{Y}_p \end{pmatrix},$$ where $\boldsymbol{z}_{1,p}$ is $(1 \times q)$ and $\boldsymbol{Z}_{2,p}$ is $(n-1 \times q)$ with

$$\boldsymbol{Z}_{2,p} = \begin{pmatrix} \boldsymbol{z}_{22,p} \\ \vdots \\ \boldsymbol{z}_{2n,p} \end{pmatrix},$$ where each component $\boldsymbol{z}_{2m,p}, m = 2, \ldots, n$ is $(1 \times q)$ dimensional vector.

$$\boldsymbol{C}'\boldsymbol{T}_p = \begin{pmatrix} \boldsymbol{u}_{1,p} \\ \boldsymbol{U}_{2,p} \end{pmatrix} = \begin{pmatrix} \frac{\boldsymbol{j}_n'}{\sqrt{n}}\boldsymbol{T}_p \\ \begin{pmatrix} \boldsymbol{c}_2' \\ \vdots \\ \boldsymbol{c}_n' \end{pmatrix} \boldsymbol{T}_p \end{pmatrix}, \quad \text{for} \quad p = 1, 2, \ldots, N, \quad \text{with}$$

$$\boldsymbol{U}_{2,p} = \begin{pmatrix} \boldsymbol{u}_{22,p}' \\ \vdots \\ \boldsymbol{u}_{2n,p}' \end{pmatrix}, \quad \text{where} \quad \boldsymbol{u}_{2m,p}' = \boldsymbol{c}_m'\boldsymbol{T}_p, \quad \text{for} \quad m = 2, 3, \ldots, n.$$

Now,

$$\begin{pmatrix} \boldsymbol{z}_{1,p}' \\ \text{vec}(\boldsymbol{Z}_{2,p}') \end{pmatrix} = \boldsymbol{C}^*\boldsymbol{y}_p^* = \begin{pmatrix} \left(\frac{\boldsymbol{j}_n'}{\sqrt{n}} \otimes \boldsymbol{I}_q\right)\boldsymbol{y}_p^* \\ \left(\begin{pmatrix} \boldsymbol{c}_2' \\ \vdots \\ \boldsymbol{c}_n' \end{pmatrix} \otimes \boldsymbol{I}_q\right) \boldsymbol{y}_p^* \end{pmatrix}, \quad \text{for} \quad p = 1, 2, \ldots, N.$$

Note,

$$E\left[\left(\frac{\boldsymbol{j}_n'}{\sqrt{n}} \otimes \boldsymbol{I}_q\right)\boldsymbol{y}_p^* \right] = \left(\frac{\boldsymbol{j}_n'}{\sqrt{n}} \otimes \boldsymbol{I}_q\right) \text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}_n' + \boldsymbol{\gamma}'\boldsymbol{T}_1') = \sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_{1,p}',$$

and

$$\begin{aligned} E\left[\left(\begin{pmatrix} \boldsymbol{c}_2' \\ \vdots \\ \boldsymbol{c}_n' \end{pmatrix} \otimes \boldsymbol{I}_q\right) \boldsymbol{y}_p^*\right] &= \begin{pmatrix} \boldsymbol{c}_2' \otimes \boldsymbol{I}_q \\ \vdots \\ \boldsymbol{c}_n' \otimes \boldsymbol{I}_q \end{pmatrix} \text{vec}(\boldsymbol{\alpha}'\boldsymbol{j}_n' + \boldsymbol{\gamma}'\boldsymbol{T}_p') \\ &= \text{vec}\left[(\boldsymbol{\alpha}'\boldsymbol{j}_n' + \boldsymbol{\gamma}'\boldsymbol{T}_p')(\boldsymbol{c}_2 \ldots \boldsymbol{c}_n)\right] = \text{vec}(\boldsymbol{\gamma}'\boldsymbol{U}_{2,p}'). \end{aligned}$$

Thus, it follows

$$\begin{pmatrix} \boldsymbol{z}_{1,p}' \\ \text{vec}(\boldsymbol{Z}_{2,p}') \end{pmatrix} \sim N_{nq}\left(\begin{pmatrix} \sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_{1,p}' \\ \text{vec}(\boldsymbol{\gamma}'\boldsymbol{U}_{2,p}') \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{z}_1} & \underset{q\times(n-1)q}{\boldsymbol{0}} \\ \underset{(n-1)q\times q}{\boldsymbol{0}} & \boldsymbol{I}_{n-1} \otimes \boldsymbol{\Sigma}_{\boldsymbol{Z}_2} \end{pmatrix}\right), \quad \text{for} \quad p = 1, 2, \ldots, N,$$

where $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ are defined in Section 3.2.2. Note that $\boldsymbol{z}_{1,p}'$ is independent of $\text{vec}(\boldsymbol{Z}_{2,p}')$. Also, their marginal distributions are given by

$$\begin{aligned} \text{vec}(\boldsymbol{z}_{1,p}') &\sim N_q(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}_{1,p}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}) \\ \text{vec}(\boldsymbol{Z}_{2,p}') = \begin{pmatrix} \boldsymbol{z}_{22,p}' \\ \vdots \\ \boldsymbol{z}_{2n,p}' \end{pmatrix} &\sim N_{(n-1)q}\left(\begin{pmatrix} \boldsymbol{c}_2' \otimes \boldsymbol{I}_q \\ \vdots \\ \boldsymbol{c}_n' \otimes \boldsymbol{I}_q \end{pmatrix} \text{vec}(\boldsymbol{\alpha}'\boldsymbol{1}_n' + \boldsymbol{\gamma}'\boldsymbol{T}_p'), \boldsymbol{I}_{n-1} \otimes \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right), \quad \text{for} \quad p = 1, 2, \ldots, N. \end{aligned}$$

Therefore $\boldsymbol{z}'_{2m,p} \sim N_q\big(\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big)$ for $m = 2,3,\ldots,n$ and $p = 1,2,\ldots,N$. Also, note that $\boldsymbol{z}'_{2m,p}$ is independent of $\boldsymbol{z}'_{2m^*,p}$ for $m \neq m^*,\ m,m^* = 2,3,\ldots,n$. Also let

$$\boldsymbol{U}_{e1} = \begin{pmatrix} \boldsymbol{u}_{1,1} \\ \vdots \\ \boldsymbol{u}_{1,N} \end{pmatrix},\ \boldsymbol{U}_{e2} = \begin{pmatrix} \boldsymbol{U}_{2,1} \\ \vdots \\ \boldsymbol{U}_{2,N} \end{pmatrix},\ \boldsymbol{Z}_{e1} = \begin{pmatrix} \boldsymbol{z}_{1,1} \\ \vdots \\ \boldsymbol{z}_{1,N} \end{pmatrix},\ \text{ and }\ \boldsymbol{Z}_{e2} = \begin{pmatrix} \boldsymbol{Z}_{2,1} \\ \vdots \\ \boldsymbol{Z}_{2,N} \end{pmatrix}.$$

Hence

$$\boldsymbol{Z}'_{e1}\boldsymbol{Z}_{e1} = \sum_{p=1}^{N} \boldsymbol{z}'_{1,p}\boldsymbol{z}_{1,p} \ \text{ and }\ \boldsymbol{Z}'_{e2}\boldsymbol{Z}_{e2} = \sum_{p=1}^{N} \boldsymbol{Z}'_{2,p}\boldsymbol{Z}_{2,p} = \sum_{p=1}^{N}\sum_{m=2}^{n} \boldsymbol{z}'_{2m,p}\boldsymbol{z}_{2m,p}.$$

We now derive MLEs of the model parameters of EGLM for N multiple observations in the following section.

5.1 MLEs of transformed EGLM for multiple observations

We derive the MLEs of all model parameters $\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ for N multiple observations. The likelihood of the transformed model can be separated into two independent likelihood equations involving the densities of $\boldsymbol{z}_{1,p}$ and $\boldsymbol{Z}'_{2m,p}$ alone. Since $\boldsymbol{\alpha}$ appears in the density of $\boldsymbol{Z}_{e1}$ alone, $\boldsymbol{Z}_{e2}$ can be ignored for its estimation. Its likelihood is given by

$$L(\boldsymbol{\alpha}'|\widehat{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{Z}_1) = \prod_{p=1}^{N} (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{z}'_{1,p} - \sqrt{n}\boldsymbol{\alpha}' - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}'_{1,p} - \sqrt{n}\boldsymbol{\alpha}' - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})\right\},$$

Simplifying this, the log-likelihood of the kernel w.r.t. $\boldsymbol{\alpha}'$ is given by

$$\ln L(\boldsymbol{\alpha}'|\widehat{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{Z}_{e1}) = \sum_{p=1}^{N} \Big(\sqrt{n}\boldsymbol{z}_{1,p}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \frac{1}{2}n\boldsymbol{\alpha}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' - \sqrt{n}\boldsymbol{u}_{1,p}\widehat{\boldsymbol{\gamma}}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}'\Big).$$

Differentiating the above log-likelihood kernel w.r.t. $\boldsymbol{\alpha}'$ and equating it to $\mathbf{0}_q$ we get

$$\widehat{\boldsymbol{\alpha}}' = \frac{\displaystyle\sum_{p=1}^{N} \boldsymbol{z}'_{1,p} - \widehat{\boldsymbol{\gamma}}' \sum_{i=1}^{N} \boldsymbol{u}'_{1,p}}{N\sqrt{n}}.$$

In order to estimate $\boldsymbol{\gamma}$ we must use the full likelihood equation which is given by

$$\begin{aligned} L(\boldsymbol{\gamma}|\widehat{\boldsymbol{\alpha}}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}, \boldsymbol{Z}_{e1}, \boldsymbol{Z}_{e2}) = & \prod_{p=1}^{N} (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{z}'_{1,p} - \sqrt{n}\widehat{\boldsymbol{\alpha}}' - \boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}'_{1,p} - \sqrt{n}\widehat{\boldsymbol{\alpha}}' - \boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})\right\} \cdot \\ & \prod_{m=2}^{n} (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\boldsymbol{z}'_{2m,p} - \boldsymbol{\gamma}'\boldsymbol{u}_{2m,p})'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{z}'_{2m,p} - \boldsymbol{\gamma}'\boldsymbol{u}'_{2m,p})\right\}, \end{aligned}$$

Simplifying this, the log-likelihood of the kernel w.r.t. $\boldsymbol{\gamma}$ is given by

$$\begin{aligned}\ln L(\boldsymbol{\gamma}|\widehat{\boldsymbol{\alpha}}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}, \boldsymbol{Z}_{e1}, \boldsymbol{Z}_{e2}) &= \sum_{p=1}^{N}\Big(\mathrm{tr}(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}'_{1,p}\boldsymbol{u}_{1,p}) - \sqrt{n}\,\mathrm{tr}(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\widehat{\boldsymbol{\alpha}}'\boldsymbol{u}_{1,p}) - \frac{1}{2}\mathrm{tr}(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p})\Big)\\ &+ \sum_{p=1}^{N}\sum_{m=2}^{n}\Big(\mathrm{tr}(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{z}'_{2m,p}\boldsymbol{u}'_{2m,p}) - \frac{1}{2}\mathrm{tr}(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p}\boldsymbol{u}'_{2m,p})\Big).\end{aligned}$$

Differentiating the above log-likelihood kernel w.r.t. $\boldsymbol{\gamma}$ and equating it to $\mathbf{0}_{(r-1)\times q}$ and plugging-in the MLE of $\boldsymbol{\alpha}'$ and simplifying we get

$$\begin{aligned}&\Big[\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{z}_{1,p} - \frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{z}_{1,i}\big)}{N} + \Big(\frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{u}_{1,p}\big)}{N} - (\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p})\Big)\widehat{\gamma}\Big]\boldsymbol{\Sigma}_{\boldsymbol{Z}_1}^{-1}\\ &+\Big[(\sum_{p=1}^{N}\sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{z}_{2m,p}) - (\sum_{p=1}^{N}\sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{u}'_{2m,p})\widehat{\gamma}\Big]\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} = \mathbf{0}_{(r-1)\times q}.\end{aligned}$$

The above equation equals to $\mathbf{0}_{(r-1)\times q}$, if the terms within the brackets both equal $\mathbf{0}_{(r-1)\times q}$. Therefore, we must simultaneously solve

$$\begin{aligned}\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{z}_{1,p} - \frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{z}_{1,p}\big)}{N} + \Big(\frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{u}_{1,p}\big)}{N} - \Big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p}\Big)\Big)\widehat{\gamma} &= \mathbf{0}_{(r-1)\times q}\\ \text{and}\quad \Big(\sum_{p=1}^{N}\sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{z}_{2m,p}\Big) - \Big(\sum_{p=1}^{N}\sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{u}'_{2m,p}\Big)\widehat{\gamma} &= \mathbf{0}_{(r-1)\times q}.\end{aligned}$$

Adding the two above equations, we obtain the following single equation

$$\begin{aligned}&\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{z}_{1,p} + \Big(\sum_{p=1}^{N}\sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{z}_{2m,p}\Big) - \frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{z}_{1,p}\big)}{N}\\ &+\Big(\frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{u}_{1,p}\big)}{N} - (\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p}) - \Big(\sum_{p=1}^{N}\sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{z}_{2m,p}\Big)\Big)\widehat{\gamma} = \mathbf{0}_{(r-1)\times q}.\end{aligned}$$

Since

$$\boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p} = \sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{u}'_{2m,p}, \quad \text{and} \quad \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p} = \sum_{m=2}^{n}\boldsymbol{u}_{2m,p}\boldsymbol{z}_{2m,p},$$

the above equation reduces to

$$\begin{aligned}\Big(\frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{u}_{1,p}\big)}{N} &- (\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p} + \sum_{p=1}^{N}\boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p})\Big)\widehat{\gamma}\\ &= \frac{\big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\big)\big(\sum_{p=1}^{N}\boldsymbol{z}_{1,p}\big)}{N} - \Big(\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{z}_{1,p} + \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}\Big).\end{aligned}$$

Hence,

$$\widehat{\boldsymbol{\gamma}} = \left(\frac{\left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\right)\left(\sum_{p=1}^{N} \boldsymbol{u}_{1,p}\right)}{N} - \left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p} + \sum_{p=1}^{N} \boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p}\right) \right)^{-1} \cdot \left(\frac{\left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\right)\left(\sum_{p=1}^{N} \boldsymbol{z}_{1,p}\right)}{N} - \left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\boldsymbol{z}_{1,p} + \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}\right) \right).$$

We thus have the following theorem.

Theorem 3 *The MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for multiple observations ($N > 1$) in the model (5.12) are given by*

$$\widehat{\boldsymbol{\alpha}}' = \frac{\sum_{p=1}^{N} \boldsymbol{z}'_{1,p} - \widehat{\boldsymbol{\gamma}}' \sum_{p=1}^{N} \boldsymbol{u}'_{1,p}}{N\sqrt{n}}, \quad \textit{and} \tag{5.13}$$

$$\widehat{\boldsymbol{\gamma}} = \left(\frac{\left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\right)\left(\sum_{p=1}^{N} \boldsymbol{u}_{1,p}\right)}{N} - \left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p} + \sum_{p=1}^{N} \boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p}\right) \right)^{-1} \cdot \left(\frac{\left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\right)\left(\sum_{p=1}^{N} \boldsymbol{z}_{1,p}\right)}{N} - \left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}\boldsymbol{z}_{1,p} + \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}\right) \right). \tag{5.14}$$

Corollary 1 *For $N = 1$, the estimates $\widehat{\boldsymbol{\alpha}}'$ and $\widehat{\boldsymbol{\gamma}}$ in (5.13) and (5.14) respectively reduce to*

$$\widehat{\boldsymbol{\alpha}}' = \frac{\boldsymbol{z}'_1 - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_1}{\sqrt{n}}, \quad \textit{and}$$

$$\widehat{\boldsymbol{\gamma}} = (\boldsymbol{U}'_2\boldsymbol{U}_2)^{-1}\boldsymbol{U}'_2\boldsymbol{Z}_2.$$

Proof 3 *Proof is straightforward.*

As expected, we see these estimates are exactly same as the pseudo-MLEs in Theorem 1, and also the least square estimates as obtained by Arnold (1979). We thus see that our extended model for N independent doubly multivariate observations truly generalizes the Arnold's (1979) model for one doubly multivariate observation.

Remark 2 *Note that in the case of replicated observations $\widehat{\boldsymbol{\alpha}}$ depends only on $\boldsymbol{Z}_{e1}$, whereas $\widehat{\boldsymbol{\gamma}}$ depends on both $\boldsymbol{Z}_{e1}$ ($\boldsymbol{z}_1$) and $\boldsymbol{Z}_{e2}$ ($\boldsymbol{Z}_2$). On the other hand, in the case of $N = 1$, $\widehat{\boldsymbol{\alpha}}$ depends only on $\boldsymbol{Z}_{e1}$, $\widehat{\boldsymbol{\gamma}}$ depends only on $\boldsymbol{Z}_{e2}$.*

We will now find the MLEs of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ for multiple observations. Note that $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ only appears in the density involving $\boldsymbol{Z}_{e1}$; hence the terms from $\boldsymbol{Z}_{e2}$ can be ignored in the likelihood. So, the likelihood of the kernel w.r.t. $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is given by

$$L(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\boldsymbol{\alpha}',\boldsymbol{\gamma},\boldsymbol{Z}_{e1}) = \prod_{p=1}^{N}(2\pi)^{-\frac{q}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}(\boldsymbol{z}'_{1,p}-\sqrt{n}\boldsymbol{\alpha}'-\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})'\boldsymbol{\Sigma}^{-1}_{\boldsymbol{z}_1}(\boldsymbol{z}'_{1,p}-\sqrt{n}\boldsymbol{\alpha}'-\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})\right\}.$$

Simplifying this, the log-likelihood of the kernel $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is given by

$$\ln L(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\boldsymbol{\alpha}',\boldsymbol{\gamma},\boldsymbol{Z}_{e1}) = \frac{N}{2}\ln(|\boldsymbol{\Sigma}^{-1}_{\boldsymbol{z}_1}|) - \frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{z}_1}\sum_{p=1}^{N}(\boldsymbol{z}'_{1,p}-\sqrt{n}\boldsymbol{\alpha}'-\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})(\boldsymbol{z}'_{1,p}-\sqrt{n}\boldsymbol{\alpha}'-\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})'\Big).$$

Differentiating the above log-likelihood w.r.t. $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and equating it to zero and simplifying we get

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1} &= \frac{\sum_{p=1}^{N}(\boldsymbol{z}'_{1,p}-\sqrt{n}\widehat{\boldsymbol{\alpha}}'-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,i})(\boldsymbol{z}'_{1,p}-\sqrt{n}\widehat{\boldsymbol{\alpha}}'-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})'}{N}\\ &= \frac{(\boldsymbol{Z}_{e1}-\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}}-\boldsymbol{U}_{e1}\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_{e1}-\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}}-\boldsymbol{U}_{e1}\widehat{\boldsymbol{\gamma}})}{N}.\end{aligned}$$

Again, note that $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ only appears in the density involve $\boldsymbol{Z}_{e2}$, hence the terms from $\boldsymbol{Z}_{e1}$ can be ignored in the likelihood.

$$L(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{\alpha}',\boldsymbol{\gamma},\boldsymbol{Z}_{e2}) = \prod_{p=1}^{N}\prod_{m=2}^{n}(2\pi)^{-\frac{q}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}(\boldsymbol{z}'_{2m,p}-\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p})'\boldsymbol{\Sigma}^{-1}_{\boldsymbol{Z}_2}(\boldsymbol{z}'_{2m,p}-\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p})\Big\}.$$

Simplifying this, the log-likelihood of the kernel w.r.t. $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ is given by

$$\ln L(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{\alpha}',\boldsymbol{\gamma},\boldsymbol{Z}_{e2}) = \frac{N(n-1)}{2}\ln(|\boldsymbol{\Sigma}^{-1}_{\boldsymbol{Z}_2}|) - \frac{1}{2}\mathrm{tr}\Big(\boldsymbol{\Sigma}^{-1}_{\boldsymbol{Z}_2}\sum_{p=1}^{N}\sum_{m=2}^{n}(\boldsymbol{z}'_{2m,p}-\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p})(\boldsymbol{z}'_{2m,p}-\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p})'\Big).$$

Differentiating the above log-likelihood w.r.t. $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ and equating it to zero and simplifying we get

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} &= \frac{\sum_{p=1}^{N}\sum_{m=2}^{n}(\boldsymbol{z}'_{2m,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2m,p})(\boldsymbol{z}'_{2m,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2m,p})'}{N(n-1)}\\ &= \frac{(\boldsymbol{Z}_{e2}-\boldsymbol{U}_{e2}\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_{e2}-\boldsymbol{U}_{e2}\widehat{\boldsymbol{\gamma}})}{N(n-1)}.\end{aligned}$$

We thus have the following theorem

Theorem 4 *The MLEs of* $\boldsymbol{\Sigma}_{\boldsymbol{Z}_1}$ *and* $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ *for* N *multiple observations in the model (5.12) are given by*

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1} = \frac{(\boldsymbol{Z}_{e1}-\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}}-\boldsymbol{U}_{e1}\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_{e1}-\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}}-\boldsymbol{U}_{e1}\widehat{\boldsymbol{\gamma}})}{N}, \quad \textit{and} \tag{5.15}$$

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} = \frac{(\boldsymbol{Z}_{e2}-\boldsymbol{U}_{e2}\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_{e2}-\boldsymbol{U}_{e2}\widehat{\boldsymbol{\gamma}})}{N(n-1)}. \tag{5.16}$$

Corollary 2 *For $N = 1$ the estimates $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}$ and $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}$ in (5.15) and (5.16) reduce to*

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1} = \underset{q\times q}{\mathbf{0}} \quad \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} \;=\; \frac{(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\widehat{\boldsymbol{\gamma}})}{(n-1)} \quad \text{respectively.}$$

Proof 4 *For $N = 1$ the estimates $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}$ in (5.15) reduces to*

$$\begin{aligned}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1} &= \frac{(\boldsymbol{z}_1 - 1\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}} - \boldsymbol{u}_1\widehat{\boldsymbol{\gamma}})'(\boldsymbol{z}_1 - 1\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}} - \boldsymbol{u}_1\widehat{\boldsymbol{\gamma}})}{N}\\ &= \frac{(\boldsymbol{z}_1 - (\boldsymbol{z}_1 - \boldsymbol{u}_1\widehat{\boldsymbol{\gamma}}) - \boldsymbol{u}_1\widehat{\boldsymbol{\gamma}})'(\boldsymbol{z}_1 - (\boldsymbol{z}_1 - \boldsymbol{u}_1\widehat{\boldsymbol{\gamma}}) - \boldsymbol{u}_1\widehat{\boldsymbol{\gamma}})}{N} = \underset{q\times q}{\mathbf{0}}.\end{aligned}$$

This is the true MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ when $N = 1$ although it is nonsensical. This is the reason why there are only pseudo-MLEs when N = 1, since the likelihood increases without limit as $\boldsymbol{\Sigma}_{\boldsymbol{z}_1} \to \underset{q\times q}{\mathbf{0}}$ (see Theorem 2). The proof of the rest of the Corollary 2 is straightforward. As expected these estimates are same as the pseudo-MLEs in Theorem 2, and also the least square estimates as obtained by Arnold (1979). We thus see that our extended model for N independent doubly multivariate observations truly generalizes the Arnold's (1979) model for one doubly multivariate observation.

Next we express the MLEs in terms of the original data, so that any statistical practitioner can easily compute the estimates of the model parameters for their datasets and carry out the required tests of hypotheses (described in Section 8) for the intercept and slope parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ to draw some conclusions about their datasets.

5.2 MLEs of the model parameters in terms of original data

Let

$$\underset{Nn\times q}{\boldsymbol{Y}^*} = \begin{pmatrix} \underset{n\times q}{\boldsymbol{Y}_1} \\ \vdots \\ \underset{n\times q}{\boldsymbol{Y}_N} \end{pmatrix} \quad \text{and} \quad \underset{Nn\times(r-1)}{\boldsymbol{T}^*} = \begin{pmatrix} \underset{n\times(r-1)}{\boldsymbol{T}_1} \\ \vdots \\ \underset{n\times(r-1)}{\boldsymbol{T}_N} \end{pmatrix}.$$

Since

$$\sum_{i=1}^{N} \boldsymbol{T}'_p = [\boldsymbol{T}'_1, \ldots, \boldsymbol{T}'_N] \begin{pmatrix} \boldsymbol{I}_n \\ \vdots \\ \boldsymbol{I}_n \end{pmatrix} = (\boldsymbol{T}^*)'(\boldsymbol{J}_n \otimes \boldsymbol{I}_n),$$

from (5.14) we have

$$\begin{aligned}\widehat{\boldsymbol{\gamma}} &= \Big(\frac{\big(\sum_{p=1}^{N}\boldsymbol{T}'_p\big)\boldsymbol{J}_n\big(\sum_{p=1}^{N}\boldsymbol{T}_p\big)}{Nn} - \sum_{p=1}^{N}\boldsymbol{T}'_p\boldsymbol{T}_p\Big)^{-1}\Big(\frac{\big(\sum_{p=1}^{N}\boldsymbol{T}'_p\big)\boldsymbol{J}_n\big(\sum_{p=1}^{N}\boldsymbol{Y}_p\big)}{Nn} - \sum_{p=1}^{N}\boldsymbol{T}'_p\boldsymbol{Y}_p\Big)\\ &= \Big(\frac{(\boldsymbol{T}^*)'(\boldsymbol{j}_N\otimes\boldsymbol{I}_n)(\boldsymbol{j}_n\boldsymbol{j}'_N \;\otimes \boldsymbol{j}'_n)\boldsymbol{T}^*}{Nn} - (\boldsymbol{T}^*)'\boldsymbol{T}^*\Big)^{-1}\Big(\frac{(\boldsymbol{T}^*)'(\boldsymbol{j}_N\otimes\boldsymbol{I}_n)(\boldsymbol{j}_n\boldsymbol{j}'_N\otimes\boldsymbol{j}'_n)\boldsymbol{Y}^*}{Nn} - (\boldsymbol{T}^*)'\boldsymbol{Y}^*\Big).\end{aligned}$$

And, in terms of the original data, the MLE of $\boldsymbol{\alpha}$ is given by

$$\begin{aligned}\widehat{\boldsymbol{\alpha}}' &= \boldsymbol{j}_n'\Bigg(\frac{\sum_{p=1}^{N}\boldsymbol{Y}_p - \Big(\sum_{p=1}^{N}\boldsymbol{T}_p\Big)\widehat{\boldsymbol{\gamma}}}{Nn}\Bigg)\\ &= \boldsymbol{j}_n'(\boldsymbol{j}_N'\otimes\boldsymbol{I}_n)\Big(\frac{\boldsymbol{Y}^* - \boldsymbol{T}^*\widehat{\boldsymbol{\gamma}}}{Nn}\Big).\end{aligned}$$

Now note that,

$$\underset{N\times q}{\boldsymbol{Z}_{e1}} = \begin{pmatrix}\underset{1\times q}{\boldsymbol{z}_{1,1}}\\ \vdots\\ \underset{1\times q}{\boldsymbol{z}_{1,N}}\end{pmatrix} = \begin{pmatrix}\frac{\boldsymbol{j}_n'}{\sqrt{n}}\underset{n\times q}{\boldsymbol{Y}_1}\\ \vdots\\ \frac{\boldsymbol{j}_n'}{\sqrt{n}}\underset{n\times q}{\boldsymbol{Y}_N}\end{pmatrix} = \Big(\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}_n'}{\sqrt{n}}\Big)\boldsymbol{Y}^* \text{ and}$$

$$\underset{N\times(r-1)}{\boldsymbol{U}_1} = \begin{pmatrix}\underset{1\times(r-1)}{\boldsymbol{u}_{1,1}}\\ \vdots\\ \underset{1\times(r-1)}{\boldsymbol{u}_{1,N}}\end{pmatrix} = \begin{pmatrix}\frac{\boldsymbol{j}_n'}{\sqrt{n}}\underset{n\times(r-1)}{\boldsymbol{T}_1}\\ \vdots\\ \frac{\boldsymbol{j}_n'}{\sqrt{n}}\underset{n\times(r-1)}{\boldsymbol{T}_N}\end{pmatrix} = \Big(\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}_n'}{\sqrt{n}}\Big)\boldsymbol{T}^*.$$

Therefore, the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ can be expressed in terms of the original data as follows

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1} = \frac{\Big((\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}_n'}{\sqrt{n}})[\boldsymbol{Y}^* - \boldsymbol{T}^*\widehat{\boldsymbol{\gamma}}] - (\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}})\Big)'\Big((\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}_n'}{\sqrt{n}})[\boldsymbol{Y}^* - \boldsymbol{T}^*\widehat{\boldsymbol{\gamma}}] - (\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}})\Big)}{N}.$$

To express the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ in terms of the original data, we first consider the direct sum (not to be confused with Kronecker sum) notation

$$\oplus_{m=2}^{n}\boldsymbol{c}_m' = \begin{pmatrix}\boldsymbol{c}_2' & \cdots & \underset{1\times q}{\boldsymbol{0}}\\ \vdots & \ddots & \vdots\\ \underset{1\times q}{\boldsymbol{0}} & \cdots & \boldsymbol{c}_n'\end{pmatrix}.$$

Therefore,

$$\boldsymbol{Z}_{e2} = \begin{pmatrix}\boldsymbol{Z}_{2,1}\\ \vdots\\ \boldsymbol{Z}_{2,N}\end{pmatrix} = \begin{pmatrix}\begin{pmatrix}\boldsymbol{z}_{22,1}\\ \vdots\\ \boldsymbol{z}_{2n,1}\end{pmatrix}\\ \vdots\\ \begin{pmatrix}\boldsymbol{z}_{22,N}\\ \vdots\\ \boldsymbol{z}_{2n,N}\end{pmatrix}\end{pmatrix} = \begin{pmatrix}\begin{pmatrix}\boldsymbol{c}_2'\boldsymbol{Y}_1\\ \vdots\\ \boldsymbol{c}_n'\boldsymbol{Y}_1\end{pmatrix}\\ \vdots\\ \begin{pmatrix}\boldsymbol{c}_2'\boldsymbol{Y}_N\\ \vdots\\ \boldsymbol{c}_n'\boldsymbol{Y}_N\end{pmatrix}\end{pmatrix} = \begin{pmatrix}(\oplus_{m=2}^{n}\boldsymbol{c}_m')(\boldsymbol{j}_{n-1}\otimes\boldsymbol{Y}_1)\\ \vdots\\ (\oplus_{m=2}^{n}\boldsymbol{c}_m')(\boldsymbol{j}_{n-1}\otimes\boldsymbol{Y}_N)\end{pmatrix} \text{ and}$$

$$\boldsymbol{U}_{e2} = \begin{pmatrix} \boldsymbol{U}_{2,1} \\ \vdots \\ \boldsymbol{U}_{2,N} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \boldsymbol{u}_{22,1} \\ \vdots \\ \boldsymbol{u}_{2n,1} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \boldsymbol{u}_{22,N} \\ \vdots \\ \boldsymbol{u}_{2n,N} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \boldsymbol{c}_2'\boldsymbol{T}_1 \\ \vdots \\ \boldsymbol{c}_n'\boldsymbol{T}_1 \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \boldsymbol{c}_2'\boldsymbol{T}_N \\ \vdots \\ \boldsymbol{c}_n'\boldsymbol{T}_N \end{pmatrix} \end{pmatrix} = \begin{pmatrix} (\oplus_{m=2}^n \boldsymbol{c}_i')(\boldsymbol{j}_{n-1} \otimes \boldsymbol{T}_1) \\ \vdots \\ (\oplus_{m=2}^n \boldsymbol{c}_i')(\boldsymbol{j}_{n-1} \otimes \boldsymbol{T}_N) \end{pmatrix}.$$

Hence,

$$\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} = \frac{\sum_{p=1}^N \left([(\boldsymbol{j}_{n-1}' \otimes \boldsymbol{Y}_p') - \widehat{\gamma}'(\boldsymbol{j}_{n-1}' \otimes \boldsymbol{T}_p')](\oplus_{m=2}^n \boldsymbol{c}_m \boldsymbol{c}_m')[(\boldsymbol{j}_{n-1} \otimes \boldsymbol{Y}_p) - (\boldsymbol{j}_{n-1} \otimes \boldsymbol{T}_p)\widehat{\gamma}]\right)}{N(n-1)}.$$

Finally note that

$$\boldsymbol{\Sigma}_2 = \frac{[\boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2] - [\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2]}{n} = \frac{\boldsymbol{\Sigma}_{\boldsymbol{z}_1} - \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}}{n}.$$

Hence by the invariant property of MLEs,

$$\begin{aligned} \widehat{\boldsymbol{\Sigma}}_2 &= \frac{\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1} - \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}}{n} \quad \text{and} \\ \widehat{\boldsymbol{\Sigma}} &= \boldsymbol{I}_n \otimes \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} + \boldsymbol{j}_n \otimes \widehat{\boldsymbol{\Sigma}}_2. \end{aligned}$$

6 Distribution of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for multiple observations

We will first derive the distribution of $\boldsymbol{\gamma}$ and then the distribution of $\boldsymbol{\alpha}$.

From (3.8) we find $\boldsymbol{z}_{1,p}$ is a $(1 \times q)$ vector for $p = 1, 2, \ldots, N$, where each $\boldsymbol{z}_{1,p}$ is i.i.d. with distribution

$$\boldsymbol{z}_{1,p} \sim N_{1,q}\left(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{u}_{1,p}\boldsymbol{\gamma}, 1, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\right),$$

where $\boldsymbol{\alpha}$ is a $(1 \times q)$ vector, $\boldsymbol{u}_{1,p}$ is a $(1 \times r-1)$ vector, and $\boldsymbol{\gamma}$ is a $(r-1 \times q)$ matrix. Now from (3.9) we have

$$\boldsymbol{Z}_{2,p} \sim N_{n-1,q}(\boldsymbol{U}_{2,p}\boldsymbol{\gamma}, \boldsymbol{I}_{n-1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}),$$

where $\boldsymbol{U}_{2,p}$ is a $(n-1 \times r-1)$ matrix. From (5.14) we have

$$\begin{aligned} \widehat{\gamma} &= (\boldsymbol{U}^*)^{-1} \left(\frac{\left(\sum_{p=1}^N \boldsymbol{u}_{1,p}'\right)\left(\sum_{p=1}^N \boldsymbol{z}_{1,p}\right)}{N} - \left(\sum_{p=1}^N \boldsymbol{u}_{1,p}'\boldsymbol{z}_{1,p} + \boldsymbol{U}_{2,p}'\boldsymbol{Z}_{2,p}\right)\right) \\ &= (\boldsymbol{U}^*)^{-1} \left(\sum_{p=1}^N (\overline{\boldsymbol{u}}_{1,*}' - \boldsymbol{u}_{1,p}')\boldsymbol{z}_{1,p} - \boldsymbol{U}_{2,p}'\boldsymbol{Z}_{2,p}\right), \end{aligned}$$

where

$$\boldsymbol{U}^* = \left(\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}(\overline{\boldsymbol{u}}_{1,*} - \boldsymbol{u}_{1,p}) - \boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p}\right),$$

with

$$\overline{\boldsymbol{u}}'_{1,*} = \frac{\sum_{p=1}^{N} \boldsymbol{u}'_{1,p}}{N} \quad \text{and} \quad \boldsymbol{u}'_{1,*} = \sum_{p=1}^{N} \boldsymbol{u}'_{1,p}.$$

Finally, recall that $\boldsymbol{Z}_{2,p}$ and $\boldsymbol{z}_{1,p}$ are independent $\forall\ p = 1, 2, \ldots, N$. Note that $(\overline{\boldsymbol{u}}'_{1,*} - \boldsymbol{u}'_{1,p})\boldsymbol{z}_{1,p}$ is a $(r-1 \times q)$ matrix and $\boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}$ is also a $(r-1 \times q)$ matrix. Using known results from the matrix normal distribution and letting $N > 1$ we have

$$\begin{aligned} \sum_{p=1}^{N}(\overline{\boldsymbol{u}}'_{1,*} - \boldsymbol{u}'_{1,p})\boldsymbol{z}_{1,p} &\sim N_{r-1,q}\left(\left(\overline{\boldsymbol{u}}'_{1,*}\boldsymbol{u}'_{1,*} - \sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p}\right)\boldsymbol{\gamma}, \sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\boldsymbol{u}_{1,p} - \overline{\boldsymbol{u}}'_{1,*}\boldsymbol{u}'_{1,*}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_1}\right) \\ -\sum_{p=1}^{N}\boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p} &\sim N_{r-1,q}\left(-\sum_{p=1}^{N}\boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p}\boldsymbol{\gamma}, \sum_{p=1}^{N}\boldsymbol{U}'_{2,p}\boldsymbol{U}_{2,p}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right). \end{aligned}$$

Therefore the distribution of $\widehat{\boldsymbol{\gamma}}$ is as follows:

$$\widehat{\boldsymbol{\gamma}} \sim N_{r-1,q}\left(\boldsymbol{\gamma}, -(\boldsymbol{U}^*)^{-1}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1} + \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right).$$

We will now derive the distribution of $\boldsymbol{\alpha}$. From (5.13) we have

$$\widehat{\boldsymbol{\alpha}}' = \frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p} - \widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}}.$$

Using known results from the matrix-variate normal distribution for $N > 1$ we have

$$\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}} \sim N_{q,1}\left(\boldsymbol{\alpha}' + \frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}, \frac{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}{Nn}, 1\right) \leftrightarrow N_q\left(\boldsymbol{\alpha}' + \frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}, \frac{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}{Nn}\right) \quad \text{and}$$

$$-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}} \sim N_{q,1}\left(-\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}, \frac{\boldsymbol{\Sigma}_{\boldsymbol{z}_1} + \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}}{N^2 n}, -\boldsymbol{u}_{1,*}(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*}\right) \leftrightarrow N_{q,1}\left(-\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}, -\frac{1}{N^2 n}(\boldsymbol{u}_{1,*}(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*})(\boldsymbol{\Sigma}_{\boldsymbol{z}_1} + \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})\right).$$

The distribution of $\widehat{\boldsymbol{\alpha}}'$ can be found by a simple technique given the above two functions are independent. Since they are both $p-$variate normal distributions, $\mathbf{0}$ covariance matrix implies independence. Thus, we will prove the above two functions are independent by showing the covariance between the above two function is $\mathbf{0}$ matrix as follows:

$$\text{Cov}\Big(-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}},\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}}\Big)=E\Big[\Big(-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}}+\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big)\Big(\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}}-\Big(\boldsymbol{\alpha}'+\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big)\Big)'\Big] \tag{6.17}$$

Expanding the above, and simplifying the first term in (6.17) we get

$$\begin{aligned}
E\Big[\Big(-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}}\Big)\Big(\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}}\Big)'\Big] &= -\frac{1}{N^2n}E\Big[\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}\sum_{p=1}^{N}\boldsymbol{z}_{1,p}\Big]\\
&= -\frac{1}{N^2n}E\Big[\Big(\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}(\overline{\boldsymbol{u}}_{1,*}-\boldsymbol{u}_{1,p})-\boldsymbol{Z}'_{2,p}\boldsymbol{U}_{2,p}\Big)(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*}\sum_{p=1}^{N}\boldsymbol{z}_{1,p}\Big]\\
&= -\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big[\boldsymbol{\alpha}+\frac{1}{N\sqrt{n}}\boldsymbol{u}_{1,*}\boldsymbol{\gamma}\Big].
\end{aligned}$$

Simplifying the second term in (6.17) we get

$$E\Big[\Big(-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}}\Big)\Big(-\Big(\boldsymbol{\alpha}'+\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big)\Big)'\Big]=\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big[\boldsymbol{\alpha}+\frac{1}{N\sqrt{n}}\boldsymbol{u}_{1,*}\boldsymbol{\gamma}\Big].$$

Simplifying the third term in (6.17) we get

$$E\Big[\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big(\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}}\Big)'\Big]=\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big[\boldsymbol{\alpha}+\frac{1}{N\sqrt{n}}\boldsymbol{u}_{1,*}\boldsymbol{\gamma}\Big].$$

Finally, the fourth term in (6.17) is

$$E\Big[\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big(-\Big(\boldsymbol{\alpha}'+\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big)\Big)'\Big]=-\frac{\boldsymbol{\gamma}'\boldsymbol{u}'_{1,*}}{N\sqrt{n}}\Big[\boldsymbol{\alpha}+\frac{1}{N\sqrt{n}}\boldsymbol{u}_{1,*}\boldsymbol{\gamma}\Big].$$

Adding up these four terms we get

$$\text{Cov}\Big(-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}},\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}}\Big)=\underset{q\times q}{\mathbf{0}}.$$

Thus, $-\frac{\widehat{\boldsymbol{\gamma}}'\sum_{p=1}^{N}\boldsymbol{u}'_{1,p}}{N\sqrt{n}}$ and $\frac{\sum_{p=1}^{N}\boldsymbol{z}'_{1,p}}{N\sqrt{n}}$ are independent. Therefore,

$$\widehat{\boldsymbol{\alpha}}'\sim N_q\Big(\boldsymbol{\alpha}',\frac{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}(N-\boldsymbol{u}_{1,*}(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*})-\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\boldsymbol{u}_{1,*}(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*}}{N^2n}\Big).$$

We thus have the following theorem.

Theorem 5 *The distributions of $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for multiple observations ($N>1$) in the model (5.12) are given by*

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}' &\sim N_q\Big(\boldsymbol{\alpha}',\frac{\boldsymbol{\Sigma}_{\boldsymbol{Z}_1}(N-\boldsymbol{u}_{1,*}(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*})-\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\boldsymbol{u}_{1,*}(\boldsymbol{U}^*)^{-1}\boldsymbol{u}'_{1,*}}{N^2n}\Big)\quad \textit{and}\\
\widehat{\boldsymbol{\gamma}} &\sim N_{r-1,q}\,(\boldsymbol{\gamma},-(\boldsymbol{U}^*)^{-1},\boldsymbol{\Sigma}_{\boldsymbol{z}_1}+\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}).
\end{aligned}$$

7 Joint complete sufficient statistics of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ and $\boldsymbol{\gamma}$ for multiple observations

The parameter space of $\boldsymbol{Y}$ when $N > 1$ is a $p(p+r+1)-$dimensional space. Consider the likelihood of the $p-$th independent sample of $\boldsymbol{Y}$.

$$L\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2} | \boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) = (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2}(\boldsymbol{z}'_{1,p} - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}'_{1,p} - \sqrt{n}\boldsymbol{\alpha}' - \boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})$$

$$\prod_{m=2}^{n} (2\pi)^{-\frac{q}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}} \exp\Big\{ -\frac{1}{2}(\boldsymbol{z}'_{2m,p} - \boldsymbol{\gamma}'\boldsymbol{u}_{2m,p})'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{z}'_{2m,p} - \boldsymbol{\gamma}'\boldsymbol{u}'_{2m,p})\Big\}$$

$$= \exp\Big\{ \mathrm{tr}\big(-\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}'_{1,p}\boldsymbol{z}_{1,p}\big) + \mathrm{tr}\big((\sqrt{n}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' + \boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})\boldsymbol{z}_{1,p}\big) + \mathrm{tr}\big(-\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{Z}'_{2,p}\boldsymbol{Z}_{2,i}\big) + \mathrm{tr}(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p})$$

$$\exp\Big\{ -\frac{1}{2}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}\boldsymbol{u}_{1,p})\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}'_{1,p}) - \frac{1}{2}\sum_{m=2}^{n} \boldsymbol{u}'_{2m,p}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p}$$

$$+\frac{n-1}{2}\ln(|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}|) + \frac{1}{2}\ln(|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}|) - \frac{nq}{2}\ln(2\pi)\Big\}.$$

Observe the last equation. The first term in the first exponent is a function of the $q(q+1)/2$ non-redundant statistics given by the symmetric matrix $\boldsymbol{K}_1\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) = \boldsymbol{z}'_{1,p}\boldsymbol{z}_{1,p}$. Moreover, let $\boldsymbol{p}_1\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big) = -\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}$.

Clearly, from the second term in the first exponent, $\boldsymbol{K}_2\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) = \boldsymbol{z}_{1,p}$, which is a function of q non-redundant statistics. Moreover, let $\boldsymbol{p}_2\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big) = (\sqrt{n}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}' + \boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}'_{1,p})$.

The third term in the first exponent is a function of the $q(q+1)/2$ non-redundant statistics given by the symmetric matrix $\boldsymbol{K}_3\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) = \boldsymbol{Z}'_{2,p}\boldsymbol{Z}_{2,p}$. Moreover, let $\boldsymbol{p}_3\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big) = -\frac{1}{2}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}$.

Finally note that $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'$ is a $q \times (r-1)$ matrix and $\boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}$ is a $(r-1) \times q$ matrix. Let

$$\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}' = \begin{pmatrix} p_{4,1,1} & \dots & p_{4,1,r-1} \\ \vdots & \ddots & \vdots \\ p_{4,p,1} & \dots & p_{4,p,r-1} \end{pmatrix}, \quad \text{and} \quad \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p} = \begin{pmatrix} f_{1,1}(\boldsymbol{Z}_{2,p}) & \dots & f_{1,p}(\boldsymbol{Z}_{2,p}) \\ \vdots & \ddots & \vdots \\ f_{r-1,1}(\boldsymbol{Z}_{2,p}) & \dots & f_{r-1,p}(\boldsymbol{Z}_{2,p}) \end{pmatrix}.$$

Therefore,

$$\big(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}\big)_{kk} = \sum_{l=1}^{r-1} p_{4,k,l} f_{l,k}(\boldsymbol{Z}_{2,p}).$$

Hence, in the fourth term in the first exponent, $\mathrm{tr}\big(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}\big)$ will involve the sum of p terms in which each depends on $r-1$ functions of $\boldsymbol{Z}_{2,p}$. Since $\boldsymbol{U}'_{2,p}$ is of full rank, each $f_{l,k}(\boldsymbol{Z}_{2,p})$ is a linear independent function of $f_{s,t}(\boldsymbol{Z}_{2,p})$, where $s = l$ and $t \neq k$, $s \neq l$ and $t = k$, or $s \neq l$ and $t \neq k$.

Therefore, $\boldsymbol{K}_4\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) = \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p}$ defines $p(r-1)$ non-redundant statistics. Moreover, let $\boldsymbol{p}_4\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big) = \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'$. Putting all of this together

$$L\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2} | \boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) =$$
$$\exp\Big\{\sum_{k=1}^{4} \mathrm{tr}\big(\boldsymbol{p}_k(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})K_k(\boldsymbol{z}'_{1,i}, \mathrm{vec}(\boldsymbol{Z}'_{2,p}))\big) + \boldsymbol{q}\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big) + \boldsymbol{H}\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big)\Big\},$$

where

$$\boldsymbol{q}\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big) = -\frac{1}{2}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}\boldsymbol{u}_{1,p})\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}(\sqrt{n}\boldsymbol{\alpha}' + \boldsymbol{\gamma}'\boldsymbol{u}'_{1,p}) - \frac{1}{2}\sum_{m=2}^{n} \boldsymbol{u}'_{2m,p}\boldsymbol{\gamma}'\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_{2m,p} + \frac{n-1}{2}\ln(|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}|) + \frac{1}{2}\ln(|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}|)$$

$$\text{and} \quad \boldsymbol{H}\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big) = -\frac{nq}{2}\ln(2\pi).$$

Since the number of dimensions in the parameter space equals the number of non-redundant statistics in $\Big(\boldsymbol{K}_1\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big), \boldsymbol{K}_2\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big), \boldsymbol{K}_3\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big), \boldsymbol{K}_4\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big)\Big)$, each $\boldsymbol{K}_g (g = 1, 2, 3, 4)$ is functionally independent of one another, and the support of $\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big)$ does not depend on any of the parameters $\big(\boldsymbol{\gamma}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big)$, then we say that $\big(\boldsymbol{z}'_{1,p}, \mathrm{vec}(\boldsymbol{Z}'_{2,p})\big)$ is a regular case of the exponential class. Moreover, since we observe N samples, then $\big(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \boldsymbol{Y}_3, \boldsymbol{Y}_4\big)$ are a set of joint complete sufficient statistics for $\big(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\alpha}', \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}, \boldsymbol{\gamma}\big)$, where

$$\begin{aligned}
\boldsymbol{Y}_1 &= \sum_{p=1}^{N} \boldsymbol{z}'_{1,p}\boldsymbol{z}_{1,p} = \boldsymbol{Z}'_{e1}\boldsymbol{Z}_{e1} \\
\boldsymbol{Y}_2 &= \sum_{p=1}^{N} \boldsymbol{z}_{1,p} = \boldsymbol{z}_{1,*} \\
\boldsymbol{Y}_3 &= \sum_{p=1}^{N} \boldsymbol{Z}'_{2,p}\boldsymbol{Z}_{2,i} = \boldsymbol{Z}'_{e2}\boldsymbol{Z}_{e2} \quad \text{and} \\
\boldsymbol{Y}_4 &= \sum_{p=1}^{N} \boldsymbol{U}'_{2,p}\boldsymbol{Z}_{2,p} = \boldsymbol{U}'_{e2}\boldsymbol{Z}_{e2}.
\end{aligned}$$

8 Tests of hypotheses

In this section we derive likelihood ratio test (LRT) statistics for testing two hypotheses: first one for the slope parameter $\boldsymbol{\gamma}$ and the second one for the intercept parameter $\boldsymbol{\alpha}$.

Recall that $\boldsymbol{\gamma}$ is a $(r-1) \times q$ matrix. Now, consider partitioning $\boldsymbol{\gamma}$ as follows:

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_{(s)} \\ \gamma_{(k)} \end{bmatrix}, \quad \text{where}$$

$\boldsymbol{\gamma}_{(s)}$ is a $(r-1-k)\times q$ matrix and $\boldsymbol{\gamma}_{(k)}$ is a $k\times q$ matrix. Moreover, recall that $\boldsymbol{T}_p$ are $\big(n\times(r-1)\big)$ matrices for $p=1,2,\ldots,N$. Now consider partitioning each $\boldsymbol{T}_p$ as

$$\boldsymbol{T}_p=(\boldsymbol{T}_{p,(s)}|\boldsymbol{T}_{p,(k)}),\quad \text{with}$$

$$\boldsymbol{T}_{p,(s)}=\begin{pmatrix}\boldsymbol{t}'_{1,(s),p}\\ \vdots\\ \boldsymbol{t}'_{n,(s),p}\end{pmatrix},\quad \text{where}\ \ \boldsymbol{t}_{i,(s),p}\ \text{is a}\ (r-1-k)\times 1\ \text{vector}\ \ (i=1,2,\ldots,n),\quad \text{and}$$

$$\boldsymbol{T}_{p,(k)}=\begin{pmatrix}\boldsymbol{t}'_{1,(k),p}\\ \vdots\\ \boldsymbol{t}'_{n,(k),p}\end{pmatrix},\quad \text{where}\ \ \boldsymbol{t}_{i,(k),p}\ \text{is a}\ k\times 1\ \text{vector}\ \ (i=1,2,\ldots,n).$$

Now define

$$\boldsymbol{T}^*_{(s)}=\begin{pmatrix}\boldsymbol{t}'_{1,(s)}\\ \vdots\\ \boldsymbol{t}'_{N,(s)}\end{pmatrix}.$$

Consider the set of hypotheses:

$$H_0:\boldsymbol{\gamma}_{(k)}=\boldsymbol{0}_{k\times q}\quad \text{versus}\quad H_1:\boldsymbol{\gamma}_{(k)}\neq\boldsymbol{0}_{k\times q}. \tag{8.18}$$

One way to test these set of hypotheses is via the LRT whose statistic is defined as

$$\Lambda=\frac{\max L(H_0)}{\max L(H_1)},\quad \text{where}$$

$\max L(H_0)$ is the likelihood under H_0 and $\max L(H_1)$ is the likelihood under H_1. Since the maximum of the likelihood equation under either alternative is given by replacing the parameters in the likelihood equation with their respective MLEs, we see that

$$\begin{aligned}\max L(H_1) &= L(\widehat{\boldsymbol{\gamma}},\widehat{\boldsymbol{\alpha}}',\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1},\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}|\boldsymbol{Z}_{e1},\boldsymbol{Z}_{e2},H_1)\\ &\prod_{p=1}^{N}\Big[(2\pi)^{-\frac{q}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}(\boldsymbol{z}'_{1,p}-\sqrt{n}\widehat{\boldsymbol{\alpha}}'-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}'_{1,p}-\sqrt{n}\widehat{\boldsymbol{\alpha}}'-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})\Big\}\Big]\cdot\\ &\prod_{m=2}^{n}(2\pi)^{-\frac{q}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}(\boldsymbol{z}'_{2m,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2m,p})'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{z}'_{2m,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2m,p})\Big\}\\ &= (2\pi)^{-\frac{Nnq}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}|^{-\frac{N}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}|^{-\frac{N(n-1)}{2}}\exp\Big\{-\frac{1}{2}Nnq\Big\}.\end{aligned}$$

Since the estimators' derivations hold for any full column rank matrix of predictor variables, it is no surprise that the likelihood under H_0 $\big($i.e., $L(H_0)\big)$ has the same reduction as under H_1 but with modified estimators for $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$, which we choose to label as $\boldsymbol{\Sigma}_{\boldsymbol{z}_1.(k)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2.(k)}$ respectively.

$$L(H_0) \;=\; (2\pi)^{-\frac{Nqn}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1.(k)}|^{-\frac{N}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2.(k)}|^{-\frac{N(n-1)}{2}}\exp\Big\{-\frac{1}{2}Nnq\Big\}, \quad \text{where}$$

$$\begin{aligned}
\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_1.(k)} &= \frac{\Big((\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}_n'}{\sqrt{n}})[\boldsymbol{Y}^*-\boldsymbol{T}^*_{(s)}\widehat{\boldsymbol{\gamma}}_{(s)}]-(\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}}_{(s)})\Big)'\Big((\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}_n'}{\sqrt{n}})[\boldsymbol{Y}^*-\boldsymbol{T}^*_{(s)}\widehat{\boldsymbol{\gamma}}_{(s)}]-(\boldsymbol{j}_N\otimes\sqrt{n}\widehat{\boldsymbol{\alpha}}_{(s)})\Big)}{N}\\
\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2.(k)} &= \frac{\sum_{p=1}^{N}\Big([(\boldsymbol{j}_{n-1}'\otimes\boldsymbol{Y}_p')-\widehat{\boldsymbol{\gamma}}_{(s)}'(\boldsymbol{j}_{n-1}'\otimes\boldsymbol{T}_{p,(s)}')](\oplus_{m=2}^{n}\boldsymbol{c}_m\boldsymbol{c}_m')[(\boldsymbol{j}_{n-1}\otimes\boldsymbol{Y}_p)-(\boldsymbol{j}_{n-1}\otimes\boldsymbol{T}_{p,(s)})\widehat{\boldsymbol{\gamma}}_{(s)}]\Big)}{N(n-1)}\\
\widehat{\boldsymbol{\gamma}}_{(s)} &= \Big(\frac{(\boldsymbol{T}^*_{(s)})'(\boldsymbol{j}_N\otimes\boldsymbol{I}_n)(\boldsymbol{j}_n\boldsymbol{j}_N'\otimes\boldsymbol{j}_n')\boldsymbol{T}^*_{(s)}}{Nn}-(\boldsymbol{T}^*_{(s)})'\boldsymbol{T}^*_{(s)}\Big)^{-1}\cdot\\
&\quad\Big(\frac{(\boldsymbol{T}^*_{(s)})'(\boldsymbol{j}_N\otimes\boldsymbol{I}_n)(\boldsymbol{j}_n\boldsymbol{j}_N'\otimes\boldsymbol{j}_n')\boldsymbol{Y}^*}{Nn}-(\boldsymbol{T}^*_{(s)})'\boldsymbol{Y}^*\Big) \quad \text{and}\\
\widehat{\boldsymbol{\alpha}}_{(s)} &= \boldsymbol{j}_n'(\boldsymbol{j}_N'\otimes\boldsymbol{I}_n)\Big(\frac{\boldsymbol{Y}^*-\boldsymbol{T}^*_{(s)}\widehat{\boldsymbol{\gamma}}_{(s)}}{Nn}\Big).
\end{aligned}$$

Therefore, the LRT is given by

$$\Lambda = \frac{\max L(H_0)}{\max L(H_1)} = \left(\frac{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{z}_1.(k)}|}{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{z}_1}|}\right)^{\frac{N}{2}}\left(\frac{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{Z}_2.(k)}|}{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{Z}_2}|}\right)^{\frac{N(n-1)}{2}}.$$

Since, under H_1, we estimate the same number of parameters as H_0 with the addition of kq parameters, it follows that

$$-2\ln(\Lambda) = -N\,\ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{z}_1.(k)}|}{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{z}_1}|}\right) - N(n-1)\,\ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{Z}_2.(k)}|}{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{Z}_2}|}\right)\to\chi^2_{kq}. \tag{8.19}$$

Thus, we can define a rejection rule with approximate significance level α as:

$$\text{Reject } H_0 \text{ if } -N\,\ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{z}_1.(k)}|}{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{z}_1}|}\right) - N(n-1)\,\ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{Z}_2.(k)}|}{|\widehat{\boldsymbol{\Sigma}}^{-1}_{\boldsymbol{Z}_2}|}\right)\geq\chi^2_{kq,\alpha}, \quad \text{where}$$

$\chi^2_{kq,\alpha}$ is the upper $-\alpha$ quantile of the χ^2_{kq} distribution. We will now develop the test statistic to test the following set of hypotheses:

$$H_0 : \boldsymbol{\alpha} = \mathbf{0}_{1\times q} \quad \text{versus} \quad H_1 : \boldsymbol{\alpha} \neq \mathbf{0}_{1\times q}. \tag{8.20}$$

Since we have already found $L(H_1)$ above, we need to calculate only $L(H_0)$ for the new null hypothesis. Note that when $\boldsymbol{\alpha} = \mathbf{0}_{1\times q}$, the MLEs of $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ remain unchanged; however, the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ changes.

To differentiate between this new value, we choose to label its MLE under $L(H_0)$ as $\boldsymbol{\Sigma}_{\boldsymbol{z}_1,\boldsymbol{\alpha}}$, where

$$\begin{aligned}
\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}} &= \frac{\sum_{p=1}^{N}(\boldsymbol{z}'_{1,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})(\boldsymbol{z}'_{1,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})'}{N} \\
&= \frac{\Big((\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}'_n}{\sqrt{n}})[\boldsymbol{Y}^*-\boldsymbol{T}^*\widehat{\boldsymbol{\gamma}}]\Big)'\Big((\boldsymbol{I}_N\otimes\frac{\boldsymbol{j}'_n}{\sqrt{n}})[\boldsymbol{Y}^*-\boldsymbol{T}^*\widehat{\boldsymbol{\gamma}}]\Big)}{N} \\
&= \frac{(\boldsymbol{Y}^*-\boldsymbol{T}^*\widehat{\boldsymbol{\gamma}})'\big(\boldsymbol{I}_N\otimes\frac{\boldsymbol{J}_n}{n}\big)(\boldsymbol{Y}^*-\boldsymbol{T}^*\widehat{\boldsymbol{\gamma}})}{N}.
\end{aligned}$$

With the above value, we can evaluate the maximum of the likelihood under H_0 as follows:

$$\begin{aligned}
\max L(H_0) &= L(\widehat{\boldsymbol{\gamma}},\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}},\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}|\boldsymbol{Z}_{e1},\boldsymbol{Z}_{e2},H_0) \\
&\prod_{p=1}^{N}\Big[(2\pi)^{-\frac{q}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}}|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}(\boldsymbol{z}'_{1,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}^{-1}(\boldsymbol{z}'_{1,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}'_{1,p})\Big\}\Big]\cdot \\
&\prod_{m=2}^{n}(2\pi)^{-\frac{q}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}|^{-\frac{1}{2}}\exp\Big\{-\frac{1}{2}(\boldsymbol{z}'_{2m,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2m,p})'\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{z}'_{2m,p}-\widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_{2m,p})\Big\} \\
&= (2\pi)^{-\frac{Nqn}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}}|^{-\frac{N}{2}}|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}|^{-\frac{N(n-1)}{2}}\exp\left\{-\frac{1}{2}Nnq\right\}.
\end{aligned}$$

Therefore,

$$\Lambda=\frac{\max L(H_0)}{\max L(H_1)}=\left(\frac{|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}}^{-1}|}{|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_1}^{-1}|}\right)^{\frac{N}{2}}.$$

Since, under H_1, we estimate the same number of parameters as H_0 with the addition of q parameters, it follows that

$$-2\ln(\Lambda)=-N\ \ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}}^{-1}|}{|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_1}^{-1}|}\right)\rightarrow\chi^2_q. \tag{8.21}$$

Thus, we can define a rejection rule with approximate significance level α as:

$$\text{Reject } H_0 \text{ if } -N\ \ln\left(\frac{|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1,\boldsymbol{\alpha}}^{-1}|}{|\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_1}^{-1}|}\right)\geq\chi^2_{q,\alpha},\quad\text{where}$$

$\chi^2_{q,\alpha}$ is the upper $-\alpha$ quantile of the χ^2_q distribution.

9 Real data examples

In this section, we demonstrate the performance of our proposed extended EGLM with two real datasets from medical studies. The datasets are constructed from the first two columns (patient ID, age and gender) of Table 1 of Sperling et al. (1990) and Left and Right focus patients' information in Table 2 of Sperling et al. (1990). with identical Patient ID. These datasets are examples of doubly multivariate data with covariance structure defined by (1.1). As described in the Introduction the metabolic rates of glucose measured in mg/100g/min at 16 locations by PET scans include 8 regions of interest: the first five, frontal, sensorimotor, temporal, parietal, and occipital are known as cortical regions (Table 2, Sperling et al. (1990)), and the last three, caudate nucleus, lenticular nucleus and thalamus are known as subcortical regions (Table 3, Sperling et al. (1990)). We will only use cortical regions data in our study. Sperling et al. (1990) measured the right-sided (R) and the left-sided (L) metabolic rates in each region of interest. Clearly, the data are doubly multivariate with $q = 5$ and $n = 2$. Both the datasets have two predictor variables age and gender, so $r = 3$. The datasets consisted of 8 patients with a Left brain hemisphere focus of the epilepsy, and 8 patients with a Left brain hemisphere focus of the epilepsy respectively. So, $N = 8$, for both the datasets. Note that these two datasets are high-dimensional as $nq > N$. So, one cannot fit the existing traditional linear models to these datasets. Nevertheless, our model not only fit linear models to these datasets, but also able to perform hypotheses testing for the model parameters. Since both the datasets have two predictor variables age and gender, the model (5.12) will have the $(1 \times 5)-$dimensional intercept parameter $\boldsymbol{\alpha}$ and $(2 \times 5)-$dimensional slope parameter $\boldsymbol{\gamma}$.

To fit the proposed model (5.12) we first compute the MLEs of the model parameters. To calculate the MLEs of the matrix model parameters we first rearrange the response variables in the dataset by grouping together the right-sided metabolic rates ($i = 1$) at the frontal, sensorimotor, temporal, parietal and occipital regions and then left-sided metabolic rates ($i = 2$) at the same regions. We do it for both the datasets. The MLEs $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\gamma}}$ for the Left focus group using Equations (5.13) and 5.14 are

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} &= \begin{pmatrix} 5.889 & 6.094 & 5.823 & 5.182 & 6.484 \end{pmatrix} \\ \text{and} \quad \widehat{\boldsymbol{\gamma}} &= \begin{pmatrix} -0.030 & -0.052 & -0.038 & -0.028 & -0.050 \\ -1.534 & -1.313 & -1.514 & -1.253 & -1.514 \end{pmatrix}, \end{aligned}$$

respectively. The MLEs $\widehat{\boldsymbol{\alpha}}$ and $\widehat{\boldsymbol{\gamma}}$ for the Right focus group are

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} &= \begin{pmatrix} 7.277 & 6.409 & 4.418 & 5.529 & 5.116 \end{pmatrix} \\ \text{and} \quad \widehat{\boldsymbol{\gamma}} &= \begin{pmatrix} -0.105 & -0.078 & -0.022 & -0.063 & -0.032 \\ -0.085 & 0.061 & 0.254 & 0.181 & 0.309 \end{pmatrix}, \end{aligned}$$

respectively. The numerical results are displayed to three (rounded) decimal places. We also calculate $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{z}_1}$ and $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2}$ for both the datasets; results are not shown here. To check whether our proposed model fits the two datasets we need to test the two hypotheses proposed in Section 8. First we test whether $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ are significant using the LRT tests developed in Section 8 for the Left focus group. The calculated value of the test statistics $-2\ln(\Lambda)$ to test $\boldsymbol{\alpha}$ as in (8.20) by our proposed method turned out to be 74.8972 with 5 degrees of freedom (df). The corresponding p-value is is negligible. Therefore, the intercept is statistically significant. The calculated value of the test statistics $-2\ln(\Lambda)$ to test $\boldsymbol{\gamma}$ for full model (i.e., for both gender and age) as in (8.18) by our proposed method turned out to be 24.9966 with 10 degrees of freedom (df). The corresponding p-value is 0.0054. Therefore, we may reject the null hypothesis, i.e., at least age or gender is statistically significant. So, we will now perform post hoc tests, separately for gender and age. The test statistics $-2\ln(\Lambda)$ to test gender only is 20.1424 with 5 df. The corresponding p-value is is 0.0012. So, gender is significant by itself. The test statistics $-2\ln(\Lambda)$ to test age only is 13.4418 with 5 df. The corresponding p-value is 0.0196. So, age is significant by itself too.

We will now perform the tests for the Right focus group. The calculated value of the test statistics $-2\ln(\Lambda)$ to test $\boldsymbol{\alpha}$ as in (8.20) by our proposed method turned out to be 85.9385 with 5 df. The corresponding p-value is negligible. Therefore, the intercept is statistically significant. The calculated value of the test statistics $-2\ln(\Lambda)$ to test $\boldsymbol{\gamma}$ for full model (i.e., for both gender and age) as in (8.18) by our proposed method turned out to be 46.0546 with 10 degrees of freedom (df). The corresponding p-value is is negligible. Therefore, we may reject the null hypothesis, i.e., at least age or gender is statistically significant. Thus, we will now perform post hoc tests, separately for gender and age, as was done for the Left focus group . The test statistics $-2\ln(\Lambda)$ to test gender is 39.3512 with 5 df. The corresponding p-value is negligible. So, gender is significant by itself. The test statistics $-2\ln(\Lambda)$ to test age is 22.7788 with 5 df. The corresponding p-value is 0.0004. So, age is significant by itself too. Thus, we see that our extended model is a good fit for both the datasets.

10 Concluding remarks and the scope for future research

In this article, we study the linear model with intercept and slope parameters for doubly multivariate data with BE covariance structure. Such a structure is a realistic assumption in many datasets. We express the MLEs in terms of the original data, so that any practitioner can use our model and do his or her data analyses much easily. We also develop LRT statistics to test the intercept and slope parameters

$\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ for our proposed model. However, it is known that LRT converges to a χ^2 distribution when the sample size is large. We are currently working on other test statistics for testing $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, like Rao's score test (Filipiak et al. (2018)) and will report the results in a future correspondence. The gradient test proposed by Terrell (Terrell (2002)) is another alternative to test the hypotheses on intercept and slope parameters. Terrell's test statistic shares the same first order asymptotic properties with the three classical tests, the likelihood ratio, the Wald and the Rao's score statistics, and is very simple when compared with the same three classical tests. We will explore Terrell's method to develop a test statistic to test the null hypotheses in the future. Our proposed model is an important contribution to fit linear models for the high-dimensional doubly multivariate data.

References

[1] S.F. Arnold, Application of the theory of products of problems to certain patterned covariance matrices. Ann Stat. 1(4) (1973) 682-699.

[2] S.F. Arnold, Applications of products to the generalized compound symmetry problem. Ann Stat. 3 (1976) 227-233.

[3] S.F. Arnold, Linear models with exchangeably distributed errors. J Am Stat Assoc 74 (1979) 194–199.

[4] R.G. Craig, Autocorrelation in LANDSAT data. Proc. 13th Internat. Symp. on Remote Sensing of Environment (1979) 1517-1524.

[5] P. Dutilleul, The MLE algorithm for the matrix normal distribution. J Stat Comput Simul. 64 (1999)105-123.

[6] K. Filipiak, D. Klein, A. Roy, A Comparison of Likelihood Ratio Tests and Rao's Score Test for Three Separable Covariance Matrix Structures, Biometrical Journal 59(1) (2017) 192-215.

[7] C. Hao, Y. Liang, A. Roy, Equivalency between Vertices and Centers-coupled-with-radii Principal Component Analyses for Interval Data. Statist Probab Lett. 106 (2015) 113-120.

[8] N. Lu, D.L. Zimmerman, The likelihood ratio test for a separable covariance matrix. Statist Probab Lett. 73 (2005) 449-457.

[9] A. Roy, A note on testing of Kronecker product covariance structures for doublymultivariate data. In: Proceedings of the American Statistical Association, Statistical Computing Section; Seattle,WA. (2007) 2157–2162.

[10] A. Roy, A new classification rule for incomplete doubly multivariate data using mixed effects model with performance comparisons on the imputed data. Stat. Med. 25(10) (2006) 1715–1728.

[11] A. Roy, R. Khattree, Tests for mean and covariance structures relevant in repeated measures based discriminant analysis, J. Appl. Statist. Sci. 12 (2003) 91–104.

[12] A. Roy, R. Khattree, On implementation of a test for Kronecker product covariance structure for multivariate repeated measures data, Stat. Methodol. 2 (2005) 297–306.

[13] A. Roy, R. Khattree, Testing the hypothesis of a Kronecker product covariance matrix in multivariate repeated measures data, in: Proceedings of the 30th Annual SAS Users Group International Conference, SUGI 30, Philadelphia, 2005.

[14] A. Roy, R. Leiva, Estimating and testing a structured covariance matrix for three-level multivariate data. Commun Stat Theory Meth. 40(10) (2011) 1945–1963.

[15] SAS Institute Inc. SAS/STAT User‘s Guide Version 9.3, Cary, NC: SAS Institute Inc. 2009.

[16] M.R. Sperling, R.C. Gur, A. Alavi, R.E. Gur, S. Resnick, M.J. O'Connor, M. Reivich, Subcortical metabolic alterations in partial epilepsy. Epilepsia 31(2) (1990) 145–155.

[17] J.A. Richards, X. Jia, Remote Sensing and Digital Image Analysis: An Introduction. Berlin: Springer Verlag. 1999.

[18] M. Srivastava, T. von Rosen, D. von Rosen, Models with a Kronecker product covariance structure: Estimation and testing, Math. Methods Statist. 17 (2008) 357–370.

[19] T.H. Szatrowski, Estimation and Testing for Block Compound Symmetry and Other Patterned Covariance Matrices With Linear and Non-linear Structure. *Technical Report Number OLK NSF 107, Stanford University, Department of Statistics*, 1976.

[20] G.R. Terrell, The gradient statistic, Comput. Sci. Stat. 34 (2002) 206–215.

[21] S.S. Wilks, Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. Ann. Math. Stat. 17(3) (1946) 257–281.

[22] I. Žežula, D. Klein, A. Roy, Testing of multivariate repeated measures data with block exchangeable covariance structure. TEST. 27(2) (2018) 360-378.