Revisiting the Linear Models with Exchangeably Distributed Errors

Timothy Opheim* Anuradha Roy †

Abstract
The popularity of the classical general linear model (CGLM) is mostly due to the ease of model building and authentication. However, the CGLM is not appropriate and thus not applicable for correlated two dimensional observations. In this article we revisit Arnold's (1979) exchangeable general linear model (EGLM) for one observation, derive profiled maximum likelihood estimates (P-MLEs) of the model's parameters, and obtain its joint complete sufficient statistics. We also obtain the joint complete sufficient statistics of the "extended" model of Arnold (1979) for the case of multiple observations.

Key Words: Block exchangeable covariance structure, linear models, profiled maximum likelihood estimates

1. Introduction

Theoretical inference in statistics is mostly based on the assumption of independent random samples drawn from an infinite population. Indeed, when the random variables exhibit dependency, models that assume independency yield misleading and misguided results. Consider a digital image where contiguous pixels are assumed to be correlated. The correlation exists because sensors take a significant amount of energy from these contiguous pixels and cover a land region much larger than the size of a pixel. Likewise, correlations also exist within each pixel since by definition, a pixel is (typically) an ordered triplet of correlated red, green, and blue coordinates. From the correlation structure, a digital image can be assumed as one sample of multivariate repeated measurements: $p = 3$ intensities are repeatedly measured over n contiguous pixels. A model based on samples of these contiguous pixels must take into account these two types of correlations, such as a linear model with errors following a matrix-variate normal distribution with some structured covariance matrix.

Matrix-variate data, where p 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 an $(n \times p)-$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. Arnold (1979) developed a linear model for an $n \times p$ matrix-variate realization of one sample of multivariate repeated measurements with a block exchangeable (BE) covariance structure (defined later) for the error term, and he labelled it the exchangeable general linear model (EGLM). The 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).

The classical linear model for multivariate (vector-variate) data can be extended to the case of doubly multivariate (matrix-variate) data, $\underset{n\times p}{\boldsymbol{Y}}$. For example, the classical general

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

†Department of Management Science and Statistics, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA

linear model (CGLM) is

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

where $\boldsymbol{X}$ is an $n \times r$ design matrix whose first column is a vector of ones, $\boldsymbol{B}$ is an $r \times p$ matrix of unknown constants and $\boldsymbol{E}$ is an $n \times p$ error matrix.

When the n rows of $\boldsymbol{Y}$ are exchangeable (see e.g., Arnold 1973, 1979 and Kozioł et al. 2018), the covariance structure of vec$(\boldsymbol{E}')$ is said to be BE and is written as

$$\begin{aligned}\underset{np\times np}{\boldsymbol{\Sigma}} &= \begin{pmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_1 \end{pmatrix} \\ &= \boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2, \end{aligned} \tag{1}$$

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, $\boldsymbol{J}_n = \mathbf{1}_n\mathbf{1}_n'$, where $\mathbf{1}_n$ is the $n-$dimensional vector of ones, and 'vec' of a matrix is a linear transformation which converts a matrix into a column vector. We assume that the $p \times p$ matrix $\boldsymbol{\Sigma}_1$ is positive definite (PD, denoted by $\boldsymbol{\Sigma}_1 > 0$), and the $p \times p$ symmetric matrix $\boldsymbol{\Sigma}_2$ satisfies $\boldsymbol{\Sigma}_1 + (n-1)\boldsymbol{\Sigma}_2 > 0$ and $\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2 > 0$ in order to assure the positive definiteness of $\boldsymbol{\Sigma}$ $\big($see Lemma 2.1 in Roy and Leiva (2011)$\big)$.

The aim of this paper is to find the joint complete sufficient statistics of an "extended" EGLM of Arnold (1979) for N multiple observations. Thereby, we first find the joint complete sufficient statistics of the transformed EGLM of Arnold (1979) for one observation.

2. Preliminaries

2.1 Exponential class, Completeness, Sufficiency, and UMVUE

Suppose $\boldsymbol{x}$ is a $p-$dimensional random vector with probability density function (pdf) $f(\boldsymbol{x};\boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \boldsymbol{\Omega} \subset \mathbb{R}^k$ and $\boldsymbol{x} \in \boldsymbol{S} \subset \mathbb{R}^p$. Let $\boldsymbol{\eta}'(\boldsymbol{\theta}) = (\eta_1(\boldsymbol{\theta}), ..., \eta_m(\boldsymbol{\theta}))'$ and $\boldsymbol{T}'(\boldsymbol{x}) = (T_1(\boldsymbol{x}), ..., T_m(\boldsymbol{x}))'$. Suppose $f(\boldsymbol{x};\boldsymbol{\theta})$ is of the form

$$f(\boldsymbol{x};\boldsymbol{\theta}) = \begin{cases} \exp\Big(\boldsymbol{\eta}'(\boldsymbol{\theta})\boldsymbol{T}(\boldsymbol{x}) + h(\boldsymbol{x}) + c(\boldsymbol{\theta})\Big) & \forall\, \boldsymbol{x} \in \boldsymbol{S}, \\ 0 & \text{elsewhere.} \end{cases} \tag{2}$$

When $m = k$, we say $f(\boldsymbol{x};\boldsymbol{\theta})$ is a member of the full $k-$dimensional exponential class if it can be expressed in the form defined by (2). If in addition $\boldsymbol{S}$ does not depend on $\boldsymbol{\theta}$, the components of $\boldsymbol{T}(\boldsymbol{x})$ are linearly independent, and there exists a one-to-one transformation from the usual parametrization to the natural parametrization, then $f(\boldsymbol{x};\boldsymbol{\theta})$ is a member of the regular $k-$dimensional exponential class and one can invoke the useful result that $\boldsymbol{T}(\boldsymbol{x})$ is a joint complete sufficient statistic for $\boldsymbol{\theta}$. When $m > k$, we say $f(\boldsymbol{x};\boldsymbol{\theta})$ is a member of the curved exponential class. In this case, the minimal sufficient statistics needn't be complete.

However, in multivariate cases where the number of scalar, vector and matrix parameters exceeds the number of multivariate observations, n, or in the univariate case when $n < m$, the result from the full exponential class defines redundant sufficient statistics. For example if $x \sim N(\mu, \sigma^2)$, then $f(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\sigma}^2)$ is a member of the regular 2-dimensional exponential class. If we have a sample of size n, then the joint complete sufficient statistics are $\big(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2\big)$, but when $n = 1$ the previous result implies that the joint complete sufficient statistics are (x, x^2), thus defining an obvious redundancy. Therefore, when this occurs one should use the Fisher-Neyman factorization theorem (Keener 2010) to find joint sufficient statistics and prove (or disprove) by definition that the joint sufficient statistics are complete. The Fisher-Neyman factorization theorem states that if $f(\boldsymbol{x};\boldsymbol{\theta})$ is a pdf, then $\boldsymbol{T}(\boldsymbol{x})$ is sufficient for $\boldsymbol{\theta}$ if and only if nonnegative functions g and h can be found such that $f(\boldsymbol{x};\boldsymbol{\theta}) = h(\boldsymbol{x})g(\boldsymbol{T}(\boldsymbol{x});\boldsymbol{\theta})$. The statistic $\boldsymbol{T}(\boldsymbol{x})$ is said to be complete if for every measurable function g, $E[g(\boldsymbol{T}(\boldsymbol{x}))] = 0\,\forall\,\boldsymbol{\theta}\in\boldsymbol{\Omega}$ implies that g is the zero function almost everywhere.

Suppose $\delta = g(\boldsymbol{\theta})$ is a parameter of interest for some function g. If $T^\star = h(\boldsymbol{T})$ for some function h and $E[T^\star] = \delta$, then $T^\star$ is the unique minimum variance unbiased estimator (UMVUE) of δ given $\boldsymbol{T}$ is a complete sufficient statistic (Lehmann-Scheffé theorem, Casella and Berger 1990).

2.2 Matrix-variate Normal Distribution

Let $\boldsymbol{M}$ be an $(r\times p)$ matrix and let $\boldsymbol{R}$ and $\boldsymbol{C}$ be $(r\times r)$ and $(p\times p)$ positive definite matrices, respectively. We write $\boldsymbol{Z}\sim N_{r,p}(\boldsymbol{M},\boldsymbol{R},\boldsymbol{C})$ to mean that $\boldsymbol{Z}$ is an $(r\times p)$ random matrix with $E[\boldsymbol{Z}] = \boldsymbol{M}$, row covariance matrix $\boldsymbol{R}$, and column covariance matrix $\boldsymbol{C}$. Moreover, a useful identity between the matrix-variate normal distribution and multivariate normal distribution is:

$$\boldsymbol{Z}\sim N_{r,p}(\boldsymbol{M},\boldsymbol{R},\boldsymbol{C}) \;\Leftrightarrow\; \operatorname{vec}(\boldsymbol{Z})\sim N_{rp}\big(\operatorname{vec}(\boldsymbol{M}),\boldsymbol{C}\otimes\boldsymbol{R}\big).$$

3. The Model

Doubly multivariate data, are data where the observations in each $(n\times p)-$dimensional matrix-variate sample are doubly correlated. As mentioned in the Introduction, the analysis of doubly multivariate data needs to take into account the correlations among the measurements of p different variables as well as the correlations among measurements taken at n different locations or time points. We briefly review the CGLM in Section 3.1 and summarize some of Arnold's results for 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 p response variables and $r-1$ predictor variables associated with it. Using the notation in Arnold (1979) with modern modifications, let $\boldsymbol{Y}$ be an $(n\times p)$ random matrix of responses, $\boldsymbol{\alpha}$ be a $(p\times 1)$ vector of parameters, $\boldsymbol{T}$ be an $(n\times(r-1))$ matrix of known constants, $\boldsymbol{\gamma}$ be an $\big((r-1)\times p\big)$ matrix of parameters, and $\boldsymbol{E}$ be an $(n\times p)$ 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 one doubly multivariate observation $\boldsymbol{Y}$ can be presented as

$$\begin{aligned}
\begin{pmatrix} \underset{1\times p}{\boldsymbol{y}_1'} \\ \vdots \\ \underset{1\times p}{\boldsymbol{y}_n'} \end{pmatrix} &= \begin{bmatrix} \underset{1\times p}{\boldsymbol{\alpha}'} + \boldsymbol{t}_1'\boldsymbol{\gamma} \\ \vdots \\ \underset{1\times p}{\boldsymbol{\alpha}'} + \boldsymbol{t}_n'\boldsymbol{\gamma} \end{bmatrix} + \begin{pmatrix} \underset{1\times p}{\boldsymbol{e}_1'} \\ \vdots \\ \underset{1\times p}{\boldsymbol{e}_n'} \end{pmatrix}, \\
\underset{n\times p}{\boldsymbol{Y}} &= \boldsymbol{1}_n\boldsymbol{\alpha}' + \underset{n\times(r-1)}{\boldsymbol{T}}\boldsymbol{\gamma} + \boldsymbol{E} \qquad (3) \\
&= \underset{n\times r}{\boldsymbol{X}}\,\underset{r\times p}{\boldsymbol{B}} + \underset{n\times p}{\boldsymbol{E}},
\end{aligned}$$

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

$$\boldsymbol{X} = \begin{bmatrix}\mathbf{1}_n : \boldsymbol{T}\end{bmatrix} \quad \text{and} \quad \underset{r\times p}{\boldsymbol{B}} = \begin{pmatrix} \underset{1\times p}{\boldsymbol{\alpha}'} \\ \underset{r-1\times p}{\boldsymbol{\gamma}} \end{pmatrix},$$

with full column rank of $\boldsymbol{X}$. The CGLM makes the following assumptions about the rows of the random error matrix $\boldsymbol{E}$: $E(\boldsymbol{e}_i) = \underset{p\times 1}{\mathbf{0}}$, $E(\boldsymbol{e}_i\boldsymbol{e}_i') = \underset{p\times p}{\boldsymbol{\Sigma}_1}$, and $E(\boldsymbol{e}_i\boldsymbol{e}_{i^*}') = \underset{p\times p}{\mathbf{0}}$, for some unknown $p\times p$ symmetric PD matrix $\boldsymbol{\Sigma}_1$ where $i, i^* \in \{1, 2, \ldots, n\}$ and $i \neq i^*$. Therefore,

$$E\big[\text{vec}(\boldsymbol{E}')\big] = E\begin{pmatrix}\boldsymbol{e}_1\\ \vdots \\ \boldsymbol{e}_n\end{pmatrix} = \begin{pmatrix}\mathbf{0}_p\\ \vdots \\ \mathbf{0}_p\end{pmatrix} = \text{vec}\left(\underset{p\times n}{\boldsymbol{O}}\right) \quad \text{and} \tag{4}$$

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

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

$$\text{vec}(\boldsymbol{E}') \sim N_{np}(\text{vec}(\underset{p\times n}{\boldsymbol{O}}), \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}_1).$$

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

$$\boldsymbol{e}_i \sim N_p(\mathbf{0}_p, \boldsymbol{\Sigma}_1), \forall\, i = 1, 2, \ldots, n.$$

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

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

3.2 Exchangeable General Linear Model

The EGLM is still defined by model (3), but it makes the following assumptions about the rows of $\boldsymbol{E}$: $E(\boldsymbol{e}_i) = \underset{p\times 1}{\mathbf{0}}$, $E(\boldsymbol{e}_i\boldsymbol{e}_i') = \underset{p\times p}{\boldsymbol{\Sigma}_1}$, and $E(\boldsymbol{e}_i\boldsymbol{e}_{i^*}') = \underset{p\times p}{\boldsymbol{\Sigma}_2}$ for some unknown $p\times p$ symmetric PD matrix $\boldsymbol{\Sigma}_1$ and some unknown $p\times p$ symmetric matrix $\boldsymbol{\Sigma}_2$, such that $\text{Cov}\big[\text{vec}(\boldsymbol{E}')\big]$ is PD where i and i^* are defined as in the previous section. Collectively, these assumptions imply the expectation matrix of vec$(\boldsymbol{E}')$ remains the same as in Equation (4), but the variance-covariance matrix of vec$(\boldsymbol{E}')$ changes to

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

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

$$\text{vec}(\boldsymbol{E}') \sim N_{np}(\text{vec}(\underset{p\times n}{\boldsymbol{O}}), \boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2).$$

It is apparent from the above formula that the rows of the response matrix are dependent random variables.

3.2.1 Eigendecomposition of Block Exchangeable Covariance Structure

Let $\boldsymbol{C}^* = \boldsymbol{C}' \otimes \boldsymbol{I}_p$, where $\boldsymbol{C} = (\frac{1}{\sqrt{n}}\mathbf{1}_n \quad \underset{n\times(n-1)}{\boldsymbol{B}_n'})$ is the $n \times n$ orthogonal Helmert matrix, with $\boldsymbol{B}_n'\boldsymbol{B}_n = \boldsymbol{Q}_n = \boldsymbol{I}_n - \frac{1}{n}\boldsymbol{J}_n$ and $\boldsymbol{B}_n\boldsymbol{B}_n' = \boldsymbol{I}_{n-1}$. Thus, $\boldsymbol{C}^*$ is also an orthogonal matrix. Recall that $\text{vec}(\boldsymbol{E}'\boldsymbol{C}) = \boldsymbol{C}^*\text{vec}(\boldsymbol{E}')$. Therefore,

$$\text{Cov}\big[\text{vec}(\boldsymbol{E}'\boldsymbol{C})\big] = \boldsymbol{C}^*[\boldsymbol{I}_n \otimes (\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_2) + \boldsymbol{J}_n \otimes \boldsymbol{\Sigma}_2]\boldsymbol{C}^{*\prime} = \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{z}_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{n-1} \otimes \boldsymbol{\Sigma}_{\boldsymbol{Z}_2} \end{bmatrix},$$

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

3.2.2 The Transformed Model

Arnold (1979) used the eigendecomposition of the BE covariance structure in his favor and diagonalized the EGLM by pre-multiplying it by the orthogonal matrix $\boldsymbol{C}'$. In this way he transformed the EGLM into n independent CGLMs (principally the principal vectors), where the variance-covariance matrices of the error terms of these n CGLMs are the n eigenblocks of the BE covariance structure (1). Let

$$\boldsymbol{C}'\boldsymbol{Y} = \begin{pmatrix} \frac{1}{\sqrt{n}}\mathbf{1}_n'\boldsymbol{Y} \\ \boldsymbol{B}_n\boldsymbol{Y} \end{pmatrix} = \begin{pmatrix} \underset{1\times p}{\boldsymbol{z}_1'} \\ \underset{(n-1)\times p}{\boldsymbol{Z}_2} \end{pmatrix} \quad \text{and} \quad \boldsymbol{C}'\boldsymbol{T} = \begin{pmatrix} \frac{1}{\sqrt{n}}\mathbf{1}_n'\boldsymbol{T} \\ \boldsymbol{B}_n\boldsymbol{T} \end{pmatrix} = \begin{pmatrix} \underset{1\times(r-1)}{\boldsymbol{u}_1'} \\ \underset{(n-1)\times(r-1)}{\boldsymbol{U}_2} \end{pmatrix},$$

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

From (3) the transformed model for one doubly multivariate observation is given by

$$\boldsymbol{C}'\boldsymbol{Y} = \boldsymbol{C}'\mathbf{1}_n\boldsymbol{\alpha}' + \boldsymbol{C}'\boldsymbol{T}\boldsymbol{\gamma} + \boldsymbol{C}'\boldsymbol{E}.$$

Hence,

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

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

$$\begin{aligned} \boldsymbol{z}_1 &\sim N_p\big(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{u}_1, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}\big) \\ \text{and} \quad \text{vec}(\boldsymbol{Z}_2') &\sim N_{(n-1)p}\big(\text{vec}(\boldsymbol{\gamma}'\boldsymbol{U}_2'), \boldsymbol{I}_{n-1} \otimes \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\big). \end{aligned}$$

In the model involving $\boldsymbol{z}_1$, there is no replication; hence, the maximum likelihood estimator of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ does not exist since the maximum is attained as $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ approaches a singular, symmetric, positive semi-definite matrix (e.g., the matrix of zeroes, $\underset{p\times p}{\boldsymbol{O}}$). The model involving $\boldsymbol{Z}_2$ has $(n-1)$ replications with no intercept term; hence for $n > 2$, sensible maximum likelihood estimators may be obtained for $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ using results from the CGLM. However, for N multiple observations, i.e., when one has $N > 1$ doubly multivariate observations and $n \geq 2$, all parameters may be estimated.

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 since the likelihood can be made arbitrarily large as $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ approaches the zero matrix. Therefore, no sensible estimators exist for $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, and hence for $\boldsymbol{\Sigma}$. We derive the profiled MLEs of the EGLM parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ for fixed $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ for one observation in the following section.

4. Profiled Maximum Likelihood Estimation of the Transformed EGLM for One Observation

Theorem 1 *The P-MLEs of $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ in the EGLM (3) for fixed $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ are given by*

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}' &= \frac{\boldsymbol{z}_1' - \widehat{\boldsymbol{\gamma}}'\boldsymbol{u}_1'}{\sqrt{n}}, \\ \widehat{\boldsymbol{\gamma}} &= (\boldsymbol{U}_2'\boldsymbol{U}_2)^{-1}\boldsymbol{U}_2'\boldsymbol{Z}_2, \\ \text{and} \quad \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Z}_2} &= \frac{\boldsymbol{V}_2|_{\gamma=\widehat{\gamma}}}{n-1}, \end{aligned}$$

where $\boldsymbol{V}_2 = (\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma})'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma})$. Moreover, the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ does not exist; hence, MLEs do not exist for the parameters of the EGLM.

Proof 1 *Suppose $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ is known. Let $\boldsymbol{V}_1 = (\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)(\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)'$ and $\boldsymbol{V}_2 = (\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma})'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma})$. The differentials of these are*

$$\begin{aligned} \mathrm{d}\boldsymbol{V}_1 &= -\sqrt{n}\mathrm{d}\boldsymbol{\alpha}(\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)' - \sqrt{n}(\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)\mathrm{d}\boldsymbol{\alpha}' \\ &- \mathrm{d}\boldsymbol{\gamma}'\boldsymbol{u}_1(\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)' - (\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)\boldsymbol{u}_1'\mathrm{d}\boldsymbol{\gamma} \\ \text{and} \quad \mathrm{d}\boldsymbol{V}_2 &= -\mathrm{d}\boldsymbol{\gamma}'\boldsymbol{U}_2'(\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma}) - (\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma})'\boldsymbol{U}_2\mathrm{d}\boldsymbol{\gamma}. \end{aligned}$$

Since $\boldsymbol{z}_1$ is independent of $\boldsymbol{Z}_2$, the logarithm of the reduced likelihood of the EGLM, up to an additive constant, is

$$l_{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{z}_1,\boldsymbol{Z}_2,\boldsymbol{\Sigma}_{\boldsymbol{z}_1}) = -\frac{1}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}| - \frac{1}{2}tr(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{V}_1) - \frac{n-1}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}| - \frac{1}{2}tr(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{V}_2).$$

The differential of the above is

$$\begin{aligned} \mathrm{d}l_{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{z}_1,\boldsymbol{Z}_2,\boldsymbol{\Sigma}_{\boldsymbol{z}_1}) &= -\frac{1}{2}tr(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\mathrm{d}\boldsymbol{V}_1) - \frac{n-1}{2}tr(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\mathrm{d}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}) \\ &- \frac{1}{2}tr(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\mathrm{d}\boldsymbol{V}_2 - \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\mathrm{d}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{V}_2) \\ &= \sqrt{n}(\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\mathrm{d}\boldsymbol{\alpha} \\ &+ vec'(\boldsymbol{u}_1(\boldsymbol{z}_1 - \sqrt{n}\boldsymbol{\alpha} - \boldsymbol{\gamma}'\boldsymbol{u}_1)'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1} \\ &+ \boldsymbol{U}_2(\boldsymbol{Z}_2 - \boldsymbol{U}_2\boldsymbol{\gamma})\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1})\mathrm{d}vec(\boldsymbol{\gamma}) \\ &+ \frac{1}{2}vec'(\boldsymbol{V}_2 - (n-1)\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})\left(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1} \otimes \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\right)\boldsymbol{D}_p\mathrm{d}vech(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}), \end{aligned}$$

which implies the three following partial derivatives:

$$\begin{aligned}\frac{\partial l_{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{z}_1,\boldsymbol{Z}_2,\boldsymbol{\Sigma}_{\boldsymbol{z}_1})}{\partial\boldsymbol{\alpha}'} &= \sqrt{n}(\boldsymbol{z}_1-\sqrt{n}\boldsymbol{\alpha}-\boldsymbol{\gamma}'\boldsymbol{u}_1)'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\\ \frac{\partial l_{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{z}_1,\boldsymbol{Z}_2,\boldsymbol{\Sigma}_{\boldsymbol{z}_1})}{\partial vec'(\boldsymbol{\gamma})} &= vec'(\boldsymbol{u}_1(\boldsymbol{z}_1-\sqrt{n}\boldsymbol{\alpha}-\boldsymbol{\gamma}'\boldsymbol{u}_1)'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}+\boldsymbol{U}_2(\boldsymbol{Z}_2-\boldsymbol{U}_2\boldsymbol{\gamma})\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1})\\ \frac{\partial l_{\boldsymbol{\Sigma}_{\boldsymbol{z}_1}}(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|\boldsymbol{z}_1,\boldsymbol{Z}_2,\boldsymbol{\Sigma}_{\boldsymbol{z}_1})}{\partial vech'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})} &= \frac{1}{2}vec'(\boldsymbol{V}_2-(n-1)\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})\left(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\otimes\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\right)\boldsymbol{D}_p,\end{aligned}$$

where $\boldsymbol{D}_p$ is the duplication matrix (see Magnus and Neudecker 1986). The P-MLEs of $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ follow upon setting the partial derivatives equal to $\mathbf{0}$ and solving the system of equations. In order to find the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$, we plug the P-MLEs into the reduced likelihood and maximize the resulting likelihood, dubbed the profiled likelihood. Specifically, the profiled likelihood, up to an additive constant, is

$$l(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\boldsymbol{z}_1,\boldsymbol{Z}_2)=-\frac{1}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|.$$

Since $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}\to\boldsymbol{O}_{p\times p}\Rightarrow|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}|\to 0$, the profiled likelihood can be made arbitrarily large as $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}\to\boldsymbol{O}_{p\times p}$; hence, the MLE of $\boldsymbol{\Sigma}_{\boldsymbol{z}_1}$ does not exist.

We will now formally discuss the statistical properties of the P-MLEs of the model parameters in the following sections. We will discuss the properties of the P-MLEs of the model parameters of the transformed EGLM for one observation in Section 4.1, and the properties of the P-MLEs of the model parameters of the "extended" transformed EGLM for multiple observations in Section 4.2.

4.1 Completeness and Sufficiency of the Transformed EGLM for One Observation

Now we wish to find the joint complete sufficient statistics for the model parameters of the EGLM. From the form of the transformed model's likelihood, we can obtain the joint complete sufficient statistics of $\boldsymbol{z}_1$ and $\boldsymbol{Z}_2$ individually and conjoin them to obtain the desired statistics. We begin with finding the joint complete sufficient statistics of $\boldsymbol{Z}_2$.

In effect, $\boldsymbol{Z}_2$ has two parameters, $\boldsymbol{\theta}_{\boldsymbol{Z}_2}=(\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$. Therefore, assume $n>2$ to avoid redundant sufficient statistics. The collection of the components of these parameters in a vector lies in the parameter space $\Omega_{\boldsymbol{Z}_2}$, an open $\frac{p(p+2r-1)}{2}$−dimensional hyperrectangle. Moreover, the support of $\boldsymbol{Z}_2$, $\boldsymbol{S}_{\boldsymbol{Z}_2}$, is $\mathbb{R}^{n-1}\times\mathbb{R}^p$. The density of $\boldsymbol{Z}_2$ is

$$f_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2|\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})=(2\pi)^{\frac{(n-1)p}{2}}|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|^{-\frac{n-1}{2}}\exp\left(-\frac{1}{2}\text{tr}\left[\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}(\boldsymbol{Z}_2-\boldsymbol{U}_2\boldsymbol{\gamma})'(\boldsymbol{Z}_2-\boldsymbol{U}_2\boldsymbol{\gamma})\right]\right).$$

Using the relation between the trace operator and the vectorize and half-vectorize operators, minor simplifications show that the above density can be rewritten in exponential class form as

$$\begin{aligned}f_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2|\boldsymbol{\gamma},\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})=\exp\Bigg[&-\frac{1}{2}\text{vech}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1})\boldsymbol{D}_p'\boldsymbol{D}_p\text{vech}(\boldsymbol{Z}_2'\boldsymbol{Z}_2)+\text{vec}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}')\text{vec}(\boldsymbol{Z}_2'\boldsymbol{U}_2)\\ &+\ h_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2)+c_{\boldsymbol{Z}_2}(\boldsymbol{\theta}_{\boldsymbol{Z}_2})\Bigg],\end{aligned}$$

where $h_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2)=\frac{n-1}{2}\ln(2\pi)$ and $c_{\boldsymbol{Z}_2}(\boldsymbol{\theta}_{\boldsymbol{Z}_2})=-\frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\boldsymbol{U}_2'\boldsymbol{U}_2\boldsymbol{\gamma}\right)-\frac{n-1}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}|$. Using the notation in Equation (2), $\boldsymbol{\eta}_{\boldsymbol{Z}_2}{}'(\boldsymbol{\theta})=(\boldsymbol{\eta}_1',\boldsymbol{\eta}_2')=\left(-\frac{1}{2}\text{vech}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1})\boldsymbol{D}_p'\boldsymbol{D}_p,\text{vec}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}')\right)$

and $\boldsymbol{T}'_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2) = \left(\text{vech}'(\boldsymbol{Z}'_2\boldsymbol{Z}_2), \text{vec}'(\boldsymbol{Z}'_2\boldsymbol{U}_2)\right)$. Since $\boldsymbol{\eta}_{\boldsymbol{Z}_2}(\boldsymbol{\theta})$ is a $\frac{p(p+2r-1)}{2}-$dimensional vector, $f_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2|\gamma, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$ is a member of the full $\frac{p(p+2r-1)}{2}-$dimensional exponential class.

Moreover, the components of $\boldsymbol{T}_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2)$ can be shown to be linearly independent by contradiction. Specifically, if the components of $\boldsymbol{T}_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2)$ are linearly dependent, then there exists constants, $c_{11}, ..., c_{1p}, c_{22}, ..., c_{p-1,p}, c_{pp}$ and $d_{11}, ..., d_{1,r-1}, d_{21}, ..., d_{p,r-1}$, not all equal to zero, such that

$$\sum_{i=1}^{p}\sum_{j=i}^{p} c_{ij}\left(\sum_{k=1}^{n-1} z_{2k,i}z_{2k,j}\right) + \sum_{i=1}^{p}\sum_{j=1}^{r-1} d_{ij}\left(\sum_{k=1}^{n-1} z_{2k,i}u_{2k,j}\right) = 0$$

for all $\boldsymbol{Z}_2 \in \boldsymbol{S}_{\boldsymbol{Z}_2}$. However, the above equation defines a quadric hypersurface which is identically equal to zero for all $\boldsymbol{Z}_2$ if and only if

$$c_{11}, ..., c_{1p}, c_{22}, ..., c_{p-1,p}, c_{pp} \quad \text{and} \quad d_{11}, ..., d_{1,r-1}, d_{21}, ..., d_{p,r-1}$$

are all equal to zero. But this is a contradiction. Hence, the components of $\boldsymbol{T}_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2)$ are linearly independent.

In addition, there exists a one-to-one transformation from the usual parametrization $(\boldsymbol{\theta}_{\boldsymbol{Z}_2})$ to the natural parametrization $(\boldsymbol{\eta}_{\boldsymbol{Z}_2})$. Note, knowledge of vech $(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$ is equivalent to knowledge of $\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}$ and the like relation holds between $\boldsymbol{\gamma}$ and vec$(\boldsymbol{\gamma})$. Therefore, $\boldsymbol{\theta}_{\boldsymbol{Z}_2}$ can be obtained from $\boldsymbol{\eta}_{\boldsymbol{Z}_2}$ by noting that

$$\text{vech}\left(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right) = -2\left(\boldsymbol{D}'_p\boldsymbol{D}_p\right)^{-1}\boldsymbol{\eta}_1 \quad \text{and} \quad \text{vec}(\boldsymbol{\gamma}') = \left(\boldsymbol{I}_{r-1}\otimes\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right)\boldsymbol{\eta}_2.$$

By definition, $f_{\boldsymbol{Z}_2}(\boldsymbol{Z}_2|\gamma, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$ is a member of the regular $\frac{p(p+2r-1)}{2}-$dimensional exponential class with joint complete sufficient statistics $\left(\text{vech}'(\boldsymbol{Z}'_2\boldsymbol{Z}_2), \text{vec}'(\boldsymbol{Z}'_2\boldsymbol{U}_2)\right)$. Now we find the complete sufficient statistic(s) of the density of $\boldsymbol{z}_1$.

The support of $\boldsymbol{z}_1$, $\boldsymbol{S}_{\boldsymbol{z}_1}$, is $\mathbb{R}^p$. Moreover, let $\boldsymbol{\theta}_{\boldsymbol{z}_1} = (\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1})$. The collection of the components of these parameters in a vector lies in the parameter space $\Omega_{\boldsymbol{z}_1}$, an open $\frac{p(p+2r+1)}{2}-$dimensional hyperrectangle. Moreover, its density can be written in the form

$$\begin{aligned} f_{\boldsymbol{z}_1}(\boldsymbol{z}_1|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}) = (2\pi)^{-\frac{p}{2}}\left|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}\right|^{-\frac{1}{2}}\exp\Big[& -\frac{1}{2}\text{vech}'\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\right)\boldsymbol{D}'_p\boldsymbol{D}_p\text{vech}\left(\boldsymbol{z}_1\boldsymbol{z}'_1\right) \\ & + \sqrt{n}\text{vec}'\left(\boldsymbol{\alpha}'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\right)\boldsymbol{z}_1 + \text{vec}'\left(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\right)\text{vec}(\boldsymbol{u}_1\boldsymbol{z}'_1) - \frac{1}{2}\text{tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\gamma}'\boldsymbol{u}_1\boldsymbol{u}'_1\boldsymbol{\gamma}\right)\Big]. \end{aligned}$$

Since $\boldsymbol{z}_1\boldsymbol{z}'_1$ and $\boldsymbol{u}_1\boldsymbol{z}'_1$ are functions of $\boldsymbol{z}_1$, by the factorization theorem, $\boldsymbol{z}_1$ is a sufficient statistic. Alternatively, since sufficiency is a data reductive tool and we have but one observation, $\boldsymbol{z}_1$ is a sufficient statistic.

We now show that $\boldsymbol{z}_1$ is complete by contradiction. Suppose there exists a measurable function g not identically zero such that $E[g(\boldsymbol{z}_1)] = 0$ for all vec$(\boldsymbol{\theta}_{\boldsymbol{z}_1}) \in \Omega_{\boldsymbol{z}_1}$. Hence,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(\boldsymbol{z}_1)\exp\left(-\frac{1}{2}\boldsymbol{z}'_1\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1\right)\exp\left(\left(\sqrt{n}\boldsymbol{\alpha}+\boldsymbol{\gamma}'\boldsymbol{u}_1\right)'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1\right)\mathrm{d}\boldsymbol{z}_1 = 0.$$

However, as a function of $(\sqrt{n}\boldsymbol{\alpha}+\boldsymbol{\gamma}'\boldsymbol{u}_1)'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}$ the above is the multidimensional bilateral laplace transform of $g(\boldsymbol{z}_1)\exp\left(-\frac{1}{2}\boldsymbol{z}'_1\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1\right)$, which cannot be identically zero unless $g(\boldsymbol{z}_1)\exp\left(-\frac{1}{2}\boldsymbol{z}'_1\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{z}_1\right)$ is zero almost everywhere. Since the latter term is never equal to zero, $g(\boldsymbol{z}_1)$ must be zero almost everywhere. Therefore, $\boldsymbol{z}_1$ is a complete sufficient statistic by contradiction.

Combining the joint complete sufficient statistics of $\boldsymbol{z}_1$ and $\boldsymbol{Z}_2$ together, we conclude that the joint complete sufficient statistics for the EGLM are $\left(\boldsymbol{z}_1, \text{vech}'(\boldsymbol{Z}'_2\boldsymbol{Z}_2), \text{vec}'(\boldsymbol{Z}'_2\boldsymbol{U}_2)\right)$.

Remark 1 *The above technique is a longer proof than that presented in Arnold (1979). Therein, he made use of the fact that the CGLM is a subfamily of the EGLM and invoked a property of completeness involving mutually absolutely continuous distributions.*

4.2 Completeness and Sufficiency of the Transformed "Extended" EGLM for Multiple Observations

We now find the joint sufficient statistics for the transformed EGLM with $N\ (>1)$ independent doubly multivariate observations. We will first explicate the site-dependent covariates case and then adumbrate the site-independent covariates case.

Using the notation in Section 3 with the addition of a subscript, let

$$\boldsymbol{Z}_{1N} = (\boldsymbol{z}_{11}, \ldots, \boldsymbol{z}_{1N}),\ \boldsymbol{U}_{1N} = (\boldsymbol{u}_{11}, \ldots, \boldsymbol{u}_{1N}),\ \text{ and }\ \boldsymbol{Z} = (\boldsymbol{Z}_{1N}, \boldsymbol{Z}_{21}, \ldots, \boldsymbol{Z}_{2N}).$$

Therefore, $\boldsymbol{Z}_{1N} \sim N_{p,N}\left(\sqrt{n}\boldsymbol{\alpha}\boldsymbol{j}'_N + \boldsymbol{\gamma}'\boldsymbol{U}_{1N}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{I}_N\right)$ and $\boldsymbol{Z}_{2i} \sim N_{n-1,p}\left(\boldsymbol{U}_{2i}\boldsymbol{\gamma}, \boldsymbol{I}_{n-1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2}\right)$ for $i = 1, \ldots, N$. In effect, the model has four parameters, $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_{\boldsymbol{z}_1}, \boldsymbol{\Sigma}_{\boldsymbol{Z}_2})$. The collection of the components of these parameters in a vector lies in the parameter space Ω, an open $p(p+r+1)-$dimensional hyperrectangle. After performing the usual manipulations, the pdf of the "extended" transformed EGLM is

$$\begin{aligned} f(\boldsymbol{Z}|\boldsymbol{\theta}) \;=\; \exp\Bigg(&-\frac{1}{2}\text{vech}'\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\right)\boldsymbol{D}'_p\boldsymbol{D}_p\text{vech}\left(\boldsymbol{Z}_{1N}\boldsymbol{Z}'_{1N}\right) + \text{vec}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}\right)\boldsymbol{Z}_{1N}\boldsymbol{1}_N \\ &+ \text{vec}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\gamma}'\right)\text{vec}\left(\boldsymbol{Z}_{1N}\boldsymbol{U}'_{1N}\right) - \frac{1}{2}\text{vech}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1})\boldsymbol{D}'_p\boldsymbol{D}_p\text{vech}\left(\sum_{i=1}^{N}\boldsymbol{Z}'_{2i}\boldsymbol{Z}_{2i}\right) \\ &+ \text{vec}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}')\text{vec}\left(\sum_{i=1}^{N}\boldsymbol{Z}'_{2i}\boldsymbol{U}_{2i}\right) + h(\boldsymbol{Z}) + c(\boldsymbol{\theta})\Bigg), \end{aligned}$$

$$\begin{aligned} \text{where}\quad h(\boldsymbol{Z}) &= -\frac{Nnp}{2}\ln(2\pi) \\ \text{and}\quad c(\boldsymbol{\theta}) &= -\frac{N}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}| - \frac{N(n-1)}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}| - \frac{1}{2}\text{tr}\left(\boldsymbol{\gamma}\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1}\boldsymbol{\gamma}'\sum_{i=1}^{N}\boldsymbol{U}'_{2i}\boldsymbol{U}_{2i}\right) \\ &- \frac{1}{2}\text{tr}\left((\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{U}_{1N})'\boldsymbol{\Sigma}'_{\boldsymbol{z}_1}(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{U}_{1N})\right). \end{aligned}$$

By the factorization theorem, the joint sufficient statistics of the above density are

$$\begin{aligned} \boldsymbol{T}(\boldsymbol{Z}) = \Bigg(&\text{vech}\left(\boldsymbol{Z}_{1N}\boldsymbol{Z}'_{1N}\right), \boldsymbol{Z}_{1N}\boldsymbol{1}_N, \text{vec}\left(\boldsymbol{Z}_{1N}\boldsymbol{U}'_{1N}\right), \\ &\text{vech}\left(\sum_{i=1}^{N}\boldsymbol{Z}'_{2i}\boldsymbol{Z}_{2i}\right), \text{vec}\left(\sum_{i=1}^{N}\boldsymbol{Z}'_{2i}\boldsymbol{U}_{2i}\right)\Bigg). \end{aligned}$$

Clearly, $\boldsymbol{T}(\boldsymbol{Z})$ has dimensions $p(p+2r)$. Since this is larger than $p(p+r+1)$ when $r > 1$ (1+ covariate included in the model), the pdf of the "extended" transformed EGLM is a member of the curved exponential family. Therefore, it should not come as a surprise that $\boldsymbol{T}(\boldsymbol{Z})$ is not complete. To see this note,

$$E\left[\left(\sum_{i=1}^{N}\boldsymbol{Z}'_{2i}\boldsymbol{U}_{2i}\right)\left(\sum_{i=1}^{N}\boldsymbol{U}'_{2i}\boldsymbol{U}_{2i}\right)^{-1}\boldsymbol{U}_{1N}\boldsymbol{Q}_N\boldsymbol{U}'_{1N} - \boldsymbol{Z}_{1N}\boldsymbol{Q}_N\boldsymbol{U}'_{1N}\right] = \boldsymbol{O}_{p\times(r-1)}.$$

By definition, $\boldsymbol{T}(\boldsymbol{Z})$ is not complete since the function within the expectation is not identically zero (almost everywhere) for all values in the parameter space.

Finally, we derive the joint complete sufficient statistics of the "extended" transformed EGLM with site-independent covariates. By definition, this implies the $i-$th doubly multivariate observation has an associated $(r-1)\times 1$ vector of covariates, $\boldsymbol{x}_i$, wherefrom the

$i-$th matrix of covariates is obtained by the relation $\boldsymbol{T}_i = \mathbf{1}_n \boldsymbol{x}_i'$. Using the notation from Section 3.2.2, $\boldsymbol{u}_{1i} = \sqrt{n}\boldsymbol{x}_i$ and $\boldsymbol{U}_{2i} = \boldsymbol{O}_{(n-1)\times(r-1)}$. Plugging in these values to the pdf of the "extended" transformed EGLM shown above, we obtain the pdf of the "extended" transformed EGLM with site-independent covariates:

$$f(\boldsymbol{Z}|\boldsymbol{\theta}) = \exp\Bigg(-\frac{1}{2}\text{vech}'\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\right)\boldsymbol{D}_p'\boldsymbol{D}_p\text{vech}\left(\boldsymbol{Z}_{1N}\boldsymbol{Z}_{1N}'\right) + \text{vec}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\alpha}\right)\boldsymbol{Z}_{1N}\mathbf{1}_N$$

$$+ \text{vec}\left(\boldsymbol{\Sigma}_{\boldsymbol{z}_1}^{-1}\boldsymbol{\gamma}'\right)\text{vec}\left(\boldsymbol{Z}_{1N}\boldsymbol{U}_{1N}'\right) - \frac{1}{2}\text{vech}'(\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}^{-1})\boldsymbol{D}_p'\boldsymbol{D}_p\text{vech}\left(\sum_{i=1}^{N}\boldsymbol{Z}_{2i}'\boldsymbol{Z}_{2i}\right)$$

$$+ h(\boldsymbol{Z}) + c(\boldsymbol{\theta})\Bigg), \tag{5}$$

where $h(\boldsymbol{Z}) = -\frac{Nnp}{2}\ln(2\pi)$

and
$$c(\boldsymbol{\theta}) = -\frac{N}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{z}_1}| - \frac{N(n-1)}{2}\ln|\boldsymbol{\Sigma}_{\boldsymbol{Z}_2}| - \frac{1}{2}\text{tr}\left((\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{U}_{1N})'\boldsymbol{\Sigma}_{\boldsymbol{z}_1}'(\sqrt{n}\boldsymbol{\alpha} + \boldsymbol{\gamma}'\boldsymbol{U}_{1N})\right).$$

Using a similar proof as that developed for showing that the density of $\boldsymbol{Z}_2$ is a member of the regular exponential family, one can show that the density (5) is a member of the regular exponential family. Hence, the joint complete sufficient statistics of the "extended" transformed EGLM with site-independent covariates are

$$\boldsymbol{T}(\boldsymbol{Z}) = \left(\text{vech}\left(\boldsymbol{Z}_{1N}\boldsymbol{Z}_{1N}'\right), \boldsymbol{Z}_{1N}\mathbf{1}_N, \text{vec}\left(\boldsymbol{Z}_{1N}\boldsymbol{U}_{1N}'\right), \text{vech}\left(\sum_{i=1}^{N}\boldsymbol{Z}_{2i}'\boldsymbol{Z}_{2i}\right)\right).$$

REFERENCES

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

Arnold, S. F. (1976), "Applications of Products to the Generalized Compound Symmetry Problem," *Annals of Statistics* , 3, 227–233.

Arnold, S. F. (1979), "Linear Models with Exchangeably Distributed Errors," *Journal of the American Statistical Association*, 74, 194–199.

Casella, G., and Berger, R. L. (1990), *Statistical Inference*, Duxbury Press.

Hao, C., Liang, Y., and Roy, A. (2015), "Equivalency Between Vertices and Centers-coupled-with-radii Principal Component Analyses for Interval Data," *Statistics & Probability Letters* , 106, 113–120.

Keener, R. W. (2010), *Theoretical Statistics*, Springer Texts in Statistics, Springer.

Kozioł, A., Roy, A., Zmyślony, R., Leiva, R. and Fonseca, M. (2018), "Free-coordinate Estimation for Doubly Multivariate Data," *Linear Algebra and its Applications*, 547, 217–239.

Magnus, J. and Neudecker, H. (1986), "Symmetry, $0-1$ Matrices and Jacobians: A Review," *Econometric Theory*, 2(2), 157–190.

Roy, A., and Leiva, R. (2011), "Estimating and Testing a Structured Covariance Matrix for Three-level Multivariate Data," *Communication in Statistics- Theory and Methods*, 40(10), 1945–1963.

Szatrowski, T. H. (1976), "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*.