Department of Mathematics

Extended GMANOVA Model With a Linearly Structured Covariance Matrix

Joseph Nzabanita, Dietrich von Rosen, Martin Singull

LiTH-MAT-R--2015/07--SE

Linköpings universitet
TEKNISKA HÖGSKOLAN

Department of Mathematics
Linköping University
S-581 83 Linköping, Sweden.

Extended GMANOVA model with a linearly structured covariance matrix

Joseph Nzabanita^{1,a}, Dietrich von Rosen^{1,b}, Martin Singull¹

¹Department of Mathematics,
Linköping University,
SE-581 83 Linköping, Sweden.
E-mail: joseph.nzabanita@liu.se
E-mail: dietrich.von.rosen@liu.se
E-mail: martin.singull@liu.se

^aDepartment of Mathematics,
University of Rwanda,
PO.Box 3900 Kigali, Rwanda
E-mail: j.nzabanita@ur.ac.rw

^bDepartment of Energy and Technology,
Swedish University of Agricultural Sciences,
SE-750 07 Uppsala, Sweden.
E-mail: dietrich.von.rosen@slu.se

Abstract

In this paper we consider the extended generalized multivariate analysis of variance (GMANOVA) with a linearly structured covariance matrix. The main theme is to find explicit estimators for the mean and for the linearly structured covariance matrix. We show how to decompose the residual space, the orthogonal complement to the mean space, into $m + 1$ orthogonal subspaces and how to derive explicit estimators of the covariance matrix from the sum of squared residuals obtained by projecting observations on those subspaces. Also an explicit estimator of the mean is derived and some properties of the proposed estimators are studied.

Keywords: estimation, extended growth curve model, GMANOVA, linearly structured covariance matrix, residuals

1 Introduction

The growth curve problems are widely studied in several research areas such as medicine, econometrics, natural sciences, social sciences, etc. These problems are related to the analysis of short time series data where a characteristic of interest is measured on each unit over several time points. The classical growth curve model (GCM) by [1] also known in the statistical literature as the generalized multivariate analysis of variance (GMANOVA) has emerged as

a powerful tool to deal with such a kind of problems. The GMANOVA model was extended later on by [2] to handle different growth profiles. The resulting model was called sum of profiles model and is known in different names such as extended growth curve model (EGCM), extended GMANOVA model, etc. In [2] an iterative algorithm to obtain the maximum likelihood estimators (MLEs), which could not be obtained explicitly, was proposed. von Rosen [3] studied the model and derived explicit MLEs under the additional nested subspaces condition on the between design matrices.

In this paper we consider the extended GMANOVA model as defined in [3]. Before we define it, we give some notations that we will use throughout this paper. $\mathcal{C}(\mathbf{A})$, $r(\mathbf{A})$ and $\text{tr}(\mathbf{A})$ denote the column space, the rank and the trace of a matrix $\mathbf{A}$ respectively. For a positive definite matrix $\mathbf{S}$ and any matrix $\mathbf{A}$, $\mathbf{P}_{\mathbf{A},\mathbf{S}} = \mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{S}^{-1}$ defines the projector onto the space $\mathcal{C}_{\mathbf{S}}(\mathbf{A})$, where the subscript $\mathbf{S}$ in $\mathcal{C}_{\mathbf{S}}(\mathbf{A})$ indicates that the inner products are defined via the positive definite matrix $\mathbf{S}^{-1}$. $\mathbf{P}_{\mathbf{A},\mathbf{S}}^o = \mathbf{I} - \mathbf{P}_{\mathbf{A},\mathbf{S}} = \mathbf{P}_{\mathbf{A}^o,\mathbf{S}^{-1}}$ is a projector onto the space $\mathcal{C}(\mathbf{A})^\perp = \mathcal{C}(\mathbf{A}^o)$, where $\mathbf{A}^o$ denotes any matrix of full rank spanning the orthogonal complement to the space $\mathcal{C}(\mathbf{A})$. If $\mathbf{S} = \mathbf{I}$, we simply write $\mathbf{P}_{\mathbf{A}}$ instead of $\mathbf{P}_{\mathbf{A},\mathbf{I}}$. The symbol $\otimes$ denotes the Kronecker product of matrices or the tensor product of linear spaces. The symbol $\oplus$ denotes the direct sum of linear spaces while the symbol $\boxplus$ denotes the direct sum of tensor spaces.

Definition 1.1 *Let $\mathbf{X} : p \times n$, $\mathbf{A}_i : p \times q_i$, $\mathbf{B}_i : q_i \times k_i$, $\mathbf{C}_i : k_i \times n$, $r_1 + p \leq n$, $i = 1, 2, \dots, m$, $\mathcal{C}(\mathbf{C}'_i) \subseteq \mathcal{C}(\mathbf{C}'_{i-1})$, $i = 2, 3, \dots, m$, where $r_i = r(\mathbf{C}_i)$. The Extended Growth Curve Model is given by*

$$\mathbf{X} = \sum_{i=1}^m \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i + \mathbf{E},$$

where columns of $\mathbf{E}$ are assumed to be independently distributed as a p -variate normal distribution with mean zero and a positive definite dispersion matrix $\mathbf{\Sigma}$; i.e. $\mathbf{E} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}_n)$. The matrices $\mathbf{A}_i$ and $\mathbf{C}_i$, often called design matrices, are known matrices whereas matrices $\mathbf{B}_i$ and $\mathbf{\Sigma}$ are unknown parameter matrices.

The model in Definition 1.1 has been extensively studied by several authors and the book by Kollo and von Rosen [4] [Chapter 4] contains useful detailed information about uniqueness, estimability conditions, moments and approximative distributions of the maximum likelihood estimators. Recently other authors considered the model with slightly different conditions. For example in [5] the explicit MLEs are presented with the nested subspace conditions on the within design matrices instead. In [6, 7] the extended growth curve model without nested subspace conditions but with orthogonal design matrices is considered and generalized least-squares estimators and their properties are studied.

In most works on the extended GMANOVA model no particular attention is made on the structure of the covariance matrix. In fact there are few articles treating the problem of structured covariance matrix although it may be important in the growth curve analysis studies. For the classical growth curve model, the most studied structure are the uniform covariance structure and the serial covariance structure, see for example, [8, 9, 10]. The main

theme of this paper is to derive explicit estimators of parameters in the extended growth curve model with a linearly structured covariance matrix, which means that for $\Sigma = (\sigma_{ij})$ the only linear structure between the elements is given by $|\sigma_{ij}| = |\sigma_{kl}| \neq 0$ and there exist at least one $(i, j) \neq (k, l)$ so that $|\sigma_{ij}| = |\sigma_{kl}| \neq 0$. The examples of linear structures for the covariance matrix are the uniform structure, the compound symmetry structure, the banded structure, the Toeplitz structure, etc.

The estimation procedure that we propose will rely on the decomposition of the residual space into $m + 1$ subspaces, see Theorem 2.3, and on the study of residuals obtained from projecting observations onto those subspaces. The paper [11] was the first to propose a residual based procedure to obtain explicit estimators for an arbitrary linear structured covariance matrix in the classical growth curve model as an alternative to iterative methods. The idea was later on applied to the sum of two profiles model in [12] and our aim here is to generalize results in [12] to the extended GMANOVA model with an arbitrary number of profiles.

2 Main idea and space decomposition

In this section we give some important results from which the main idea of our discussion is derived. To start with, we recall the theorem stated and proved in [3] about the estimation of the mean structure in the extended GMANOVA model.

Theorem 2.1 *Let $\hat{B}_i$'s be the maximum likelihood estimators of B_i 's in the model as in Definition 1.1. Then*

$$P_r \sum_{i=r}^m A_i \hat{B}_i C_i = \sum_{i=r}^m (I - T_i) X C'_i (C_i C'_i)^{-} C_i,$$

where

$$\begin{aligned} P_r &= T_{r-1} T_{r-2} \times \cdots \times T_0, T_0 = I, r = 1, 2, \dots, m+1, \\ T_i &= I - P_i A_i (A'_i P'_i S_i^{-1} P_i A_i)^{-} A'_i P'_i S_i^{-1}, i = 1, 2, \dots, m, \\ S_i &= \sum_{j=1}^i K_j, i = 1, 2, \dots, m, \\ K_j &= P_j X (C'_{j-1} (C_{j-1} C'_{j-1})^{-} C_{j-1} - C'_j (C_j C'_j)^{-} C_j) X' P'_j, C_0 = I. \end{aligned}$$

A useful result is the corollary of this theorem when $r = 1$, which gives the estimated mean structure, i.e.,

$$\widehat{E[X]} = \sum_{i=1}^m A_i \hat{B}_i C_i = \sum_{i=1}^m (I - T_i) X C'_i (C_i C'_i)^{-} C_i. \quad (1)$$

Replacing T_i in (1) by its expression given in Theorem 2.1 we get

$$\widehat{E[X]} = \sum_{i=1}^m P_i A_i (A'_i P'_i S_i^{-1} P_i A_i)^{-} A'_i P'_i S_i^{-1} X C'_i (C_i C'_i)^{-} C_i,$$

or equivalently

$$\widehat{E[\mathbf{X}]} = \sum_{i=1}^m \mathbf{P}_{\mathbf{P}_i \mathbf{A}_i, \mathbf{S}_i} \mathbf{X} \mathbf{P}_{\mathbf{C}'_i}. \quad (2)$$

Noticing that the matrix $\mathbf{P}_{\mathbf{P}_i \mathbf{A}_i, \mathbf{S}_i}$ and $\mathbf{P}_{\mathbf{C}'_i}$ are projector matrices, we see that estimators of the mean structure is based on a projection of the observations on the space generated by the design matrices. Naturally, the estimators of the variance parameters are based on a projection of the observations on the residual space, that is the orthogonal complement to the design space.

If $\mathbf{\Sigma}$ would have been known, we would have from least squares theory the best linear estimator (BLUE) given by

$$\widetilde{E[\mathbf{X}]} = \sum_{i=1}^m \mathbf{P}_{\tilde{\mathbf{P}}_i \mathbf{A}_i, \mathbf{\Sigma}} \mathbf{X} \mathbf{P}_{\mathbf{C}'_i}, \quad (3)$$

where $\mathbf{S}_i$ in $\mathbf{P}_i$ is replaced with $\mathbf{\Sigma}$ to get $\tilde{\mathbf{P}}_i$. Thus, we see that in the projections, if $\mathbf{\Sigma}$ is unknown, the parameter has been replaced with $\mathbf{S}_i$'s, which according to their expressions are not maximum likelihood estimators. However, $\mathbf{S}_i$'s define consistent estimators of $\mathbf{\Sigma}$ in the sense that $n^{-1} \mathbf{S}_i \rightarrow \mathbf{\Sigma}$ in probability.

Applying the vec-operator on both sides of (3) we get

$$\text{vec}(\widetilde{E[\mathbf{X}]}) = \sum_{i=1}^m (\mathbf{P}_{\mathbf{C}'_i} \otimes \mathbf{P}_{\tilde{\mathbf{P}}_i \mathbf{A}_i, \mathbf{\Sigma}}) \text{vec} \mathbf{X}.$$

The next theorem is essential for the development of the sequel of this paper.

Theorem 2.2 *Let $\mathbf{P} = \sum_{i=1}^m \mathbf{P}_{\mathbf{C}'_i} \otimes \mathbf{P}_{\tilde{\mathbf{P}}_i \mathbf{A}_i, \mathbf{\Sigma}}$ and $\mathcal{V}_i = \mathcal{C}_{\mathbf{\Sigma}}(\tilde{\mathbf{P}}_i \mathbf{A}_i)$, $i = 1, 2, \dots, m$. Then,*

(i) *The subspaces $\mathcal{V}_i$'s are mutually orthogonal and*

$$\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_m = \mathcal{C}_{\mathbf{\Sigma}}(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_m), \quad i = 1, 2, \dots, m;$$

(ii) *The matrix $\mathbf{P}$ is a projection matrix;*

(iii) $\mathcal{C}(\mathbf{P}) = \sum_{i=1}^m \mathcal{C}(\mathbf{C}'_i) \otimes \mathcal{V}_i$;

Proof. (i) From the definition of $\tilde{\mathbf{P}}_i$'s and using Theorem 1.2.16. and Proposition 1.2.2. in [4], we have the following

$$\begin{aligned} \mathcal{V}_1 &= \mathcal{C}_{\mathbf{\Sigma}}(\tilde{\mathbf{P}}_1 \mathbf{A}_1) = \mathcal{C}_{\mathbf{\Sigma}}(\mathbf{A}_1), \\ \mathcal{V}_2 &= \mathcal{C}_{\mathbf{\Sigma}}(\tilde{\mathbf{P}}_2 \mathbf{A}_2) \\ &= \mathcal{C}_{\mathbf{\Sigma}}((\mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \mathbf{\Sigma}}) \mathbf{A}_2) \\ &= \mathcal{C}_{\mathbf{\Sigma}}(\mathbf{A}_1 : \mathbf{A}_2) \cap \mathcal{C}_{\mathbf{\Sigma}}(\mathbf{A}_1)^\perp, \end{aligned}$$

$$\begin{aligned}
\mathcal{V}_3 &= \mathcal{C}_\Sigma(\tilde{P}_3 \mathbf{A}_3) \\
&= \mathcal{C}_\Sigma((I - P_{\tilde{P}_2 \mathbf{A}_2, \Sigma}) \tilde{P}_2 \mathbf{A}_3) \\
&= \mathcal{C}_\Sigma(\tilde{P}_2(\mathbf{A}_2 : \mathbf{A}_3)) \cap \mathcal{C}_\Sigma(\tilde{P}_2 \mathbf{A}_2)^\perp, \\
&= \mathcal{C}_\Sigma((I - P_{\mathbf{A}_1, \Sigma})(\mathbf{A}_2 : \mathbf{A}_3)) \cap \mathcal{C}_\Sigma((I - P_{\mathbf{A}_1, \Sigma}) \mathbf{A}_2)^\perp, \\
&= (\mathcal{C}_\Sigma(\mathbf{A}_1)^\perp \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3)) \cap (\mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2) \cap \mathcal{C}_\Sigma(\mathbf{A}_1)^\perp)^\perp, \\
&= (\mathcal{C}_\Sigma(\mathbf{A}_1)^\perp \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3)) \cap (\mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2)^\perp + \mathcal{C}_\Sigma(\mathbf{A}_1)), \\
&= \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3) \cap (\mathcal{C}_\Sigma(\mathbf{A}_1)^\perp \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2)^\perp + \mathcal{C}_\Sigma(\mathbf{A}_1)^\perp \cap \mathcal{C}_\Sigma(\mathbf{A}_1)), \\
&= \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3) \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2)^\perp,
\end{aligned}$$

and in general

$$\mathcal{V}_i = \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_i) \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_{i-1})^\perp, \quad i = 1, 2, \dots, m, \quad \mathbf{A}_0 = \mathbf{0}.$$

This shows that the subspaces $\mathcal{V}_i$'s are mutually orthogonal. Now we prove the second assertion. Clearly

$$\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_i \subseteq \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_i).$$

Since $\mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_{i-1}) \subset \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_i)$,

$$\dim(\mathcal{V}_i) = r(\mathbf{A}_1 : \dots : \mathbf{A}_i) - r(\mathbf{A}_1 : \dots : \mathbf{A}_{i-1}),$$

so that

$$\begin{aligned}
\sum_{j=1}^i \dim(\mathcal{V}_j) &= r(\mathbf{A}_1) + r(\mathbf{A}_1 : \mathbf{A}_2) - r(\mathbf{A}_1) + \dots \\
&\quad + r(\mathbf{A}_1 : \dots : \mathbf{A}_i) - r(\mathbf{A}_1 : \dots : \mathbf{A}_{i-1}) \\
&= r(\mathbf{A}_1 : \dots : \mathbf{A}_i).
\end{aligned}$$

This proves that the dimension of $\mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \dots \oplus \mathcal{V}_i$ equals the dimension of $\mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \dots : \mathbf{A}_i)$ and hence the result follows.

(ii) We need to show that P is an idempotent matrix. Let $\mathbf{H}_i = P_{\tilde{P}_i \mathbf{A}_i, \Sigma}$ and $\mathbf{G}_i = P_{C'_i} \otimes \mathbf{H}_i$. Then, $P = \sum_{i=1}^m \mathbf{G}_i$ and

$$PP = \sum_{i=1}^m \mathbf{G}_i^2 + \sum_{i \neq j} \mathbf{G}_i \mathbf{G}_j.$$

On one hand, $\mathbf{G}_i$ is obviously an idempotent matrix as $P_{C'_i}$ and $\mathbf{H}_i$ are. On the other hand, from calculations in (i), $\mathbf{H}_i$ may be rewritten as

$$\mathbf{H}_i = P_{\mathcal{A}_i, \Sigma} - P_{\mathcal{A}_{i-1}, \Sigma},$$

where $\mathcal{A}_i = (\mathbf{A}_1 : \dots : \mathbf{A}_i)$ and thus,

$$\sum_{i=1}^m \mathbf{H}_i = P_{\mathcal{A}_m, \Sigma} \quad \text{and} \quad \sum_{i=1}^m r(\mathbf{H}_i) = r(P_{\mathcal{A}_m, \Sigma}) = r(\mathcal{A}_m).$$

So, from Lemma 3 in [13] it follows that $\mathbf{H}_i \mathbf{H}_j = \mathbf{0}$ for $i \neq j$, which in turn implies that $\mathbf{G}_i \mathbf{G}_j = \mathbf{0}$ for $i \neq j$ since $\mathbf{G}_i \mathbf{G}_j = \mathbf{P}_{\mathbf{C}'_i} \mathbf{P}_{\mathbf{C}'_j} \otimes \mathbf{H}_i \mathbf{H}_j$. Hence $\mathbf{P}^2 = \mathbf{P}$ and $\mathbf{P}$ is a projector.

(iii) With notations and calculations introduced in the proof of (ii) it is clear that $\mathcal{C}(\mathbf{P}) = \sum_{i=1}^m \mathcal{C}(\mathbf{G}_i)$. But,

$$\mathcal{C}(\mathbf{G}_i) = \mathcal{C}(\mathbf{P}_{\mathbf{C}'_i} \otimes \mathbf{H}_i) = \mathcal{C}(\mathbf{P}_{\mathbf{C}'_i}) \otimes \mathcal{C}(\mathbf{H}_i) = \mathcal{C}(\mathbf{P}_{\mathbf{C}'_i}) \otimes \mathcal{V}_i,$$

which completes the proof of Theorem 2.2. ■

We refer to the space $\mathcal{C}(\mathbf{P})$ as the mean space and it is used to estimate the mean parameters whereas $\mathcal{C}(\mathbf{P})^\perp$ is referred to as the residual space and it is used to create residuals.

When Σ is not known it should be estimated. The general idea is to use the variation in the residuals. For our purposes we decompose the residual space into $m + 1$ orthogonal subspaces and Theorem 2.3 shows how such a decomposition is made.

Theorem 2.3 *Let $\mathcal{C}(\mathbf{P})$ and $\mathcal{V}_i$ be given as in Theorem 2.2. Then*

$$\mathcal{C}(\mathbf{P})^\perp = I_1 \boxplus I_2 \boxplus \cdots \boxplus I_{m+1},$$

where

$$\begin{aligned} I_r &= \mathcal{W}_{m-r+2} \otimes (\oplus_{i=1}^{r-1} \mathcal{V}_i)^\perp, \quad r = 1, 2, \dots, m+1, \\ \mathcal{W}_r &= \mathcal{C}(\mathbf{C}'_{m-r+1}) \cap \mathcal{C}(\mathbf{C}'_{m-r+2})^\perp, \quad r = 1, \dots, m+1, \end{aligned}$$

in which by convenience $(\oplus_{i=1}^0 \mathcal{V}_i)^\perp = \emptyset^\perp = \mathcal{V}_0 = \mathbb{R}^p$, $\mathbf{C}_0 = \mathbf{I}$ and $\mathbf{C}_{m+1} = \mathbf{0}$.

Proof. On one hand, the conditions $\mathcal{C}(\mathbf{C}'_i) \subseteq \mathcal{C}(\mathbf{C}'_{i-1})$, $i = 2, 3, \dots, m$, imply that $\mathcal{C}(\mathbf{C}'_1)$ can be decomposed as a sum of orthogonal subspaces as follows:

$$\begin{aligned} \mathcal{C}(\mathbf{C}'_1) &= [\mathcal{C}(\mathbf{C}'_1) \cap \mathcal{C}(\mathbf{C}'_2)^\perp] \oplus [\mathcal{C}(\mathbf{C}'_2) \cap \mathcal{C}(\mathbf{C}'_3)^\perp] \oplus \cdots \\ &\quad \oplus [\mathcal{C}(\mathbf{C}'_{m-1}) \cap \mathcal{C}(\mathbf{C}'_m)^\perp] \oplus \mathcal{C}(\mathbf{C}'_m). \end{aligned}$$

On the other hand, by Theorem 2.2, the subspaces $\mathcal{V}_i$, $i = 1, 2, \dots, m$, are orthogonal. Hence, the result follows by letting

$$\begin{aligned} \mathcal{W}_1 &= \mathcal{C}(\mathbf{C}'_m), \quad \mathcal{W}_2 = \mathcal{C}(\mathbf{C}'_{m-1}) \cap \mathcal{C}(\mathbf{C}'_m)^\perp, \quad \dots, \\ \mathcal{W}_m &= \mathcal{C}(\mathbf{C}'_1) \cap \mathcal{C}(\mathbf{C}'_2)^\perp, \quad \mathcal{W}_{m+1} = \mathcal{C}(\mathbf{C}'_1)^\perp, \\ \mathcal{V}_{m+1} &= (\oplus_{i=1}^m \mathcal{V}_i)^\perp, \quad \mathcal{V}_0 = \oplus_{i=1}^m \mathcal{V}_i \oplus \mathcal{V}_{m+1}, \end{aligned}$$

which completes the proof. ■

The residuals obtained by projecting data to these subspaces are

$$\mathbf{H}_r = (\mathbf{I} - \sum_{i=1}^{r-1} \mathbf{P}_{\tilde{\mathbf{P}}_i \mathbf{A}_i, \Sigma}) \mathbf{X}(\mathbf{P}_{\mathbf{C}'_{r-1}} - \mathbf{P}_{\mathbf{C}'_r}), \quad (4)$$

$r = 1, 2, 3, \dots, m+1$, and here we use for convenience $\sum_{i=1}^0 \mathbf{P}_{\mathbf{P}_i \mathbf{A}_i, \mathbf{S}_i} = \mathbf{0}$.

To get more insight on what is going on, we are going to illustrate the space decomposition for $m = 3$. In this case the BLUE of the mean is

$$\widetilde{E[\mathbf{X}]} = \widetilde{\mathbf{M}}_1 + \widetilde{\mathbf{M}}_2 + \widetilde{\mathbf{M}}_3,$$

where

$$\begin{aligned}\widetilde{\mathbf{M}}_1 &= \mathbf{P}_{\mathbf{A}_1, \Sigma} \mathbf{X} \mathbf{P}_{\mathbf{C}'_1}, \\ \widetilde{\mathbf{M}}_2 &= \mathbf{P}_{\mathcal{T}_1 \mathbf{A}_2, \Sigma} \mathbf{X} \mathbf{P}_{\mathbf{C}'_2}, \quad \mathcal{T}_1 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \Sigma} = \mathbf{T}_1, \\ \widetilde{\mathbf{M}}_3 &= \mathbf{P}_{\mathcal{T}_2 \mathbf{A}_3, \Sigma} \mathbf{X} \mathbf{P}_{\mathbf{C}'_3}, \quad \mathcal{T}_2 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \Sigma} - \mathbf{P}_{\mathcal{T}_1 \mathbf{A}_2, \Sigma} = \mathbf{T}_2 \mathbf{T}_1.\end{aligned}$$

From here we see that the estimated mean is obtained by projecting observations on some subspaces. The matrices $\mathbf{P}_{\mathbf{A}_1, \Sigma}$, $\mathbf{P}_{\mathcal{T}_1 \mathbf{A}_2, \Sigma}$ and $\mathbf{P}_{\mathcal{T}_2 \mathbf{A}_3, \Sigma}$ are projectors onto the subspaces $\mathcal{V}_1 = \mathcal{C}_\Sigma(\mathbf{A}_1)$, $\mathcal{V}_2 = \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2) \cap \mathcal{C}_\Sigma(\mathbf{A}_1)^\perp$ and $\mathcal{V}_3 = \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3) \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2)^\perp$ respectively. Figure 1 shows the whole space decomposed into mean and residual subspaces.

Figure 1: Decomposition of the whole space according to the within and between individuals design matrices illustrating the mean and residual spaces: $\mathcal{V}_1 = \mathcal{C}_\Sigma(\mathbf{A}_1)$, $\mathcal{V}_2 = \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2) \cap \mathcal{C}_\Sigma(\mathbf{A}_1)^\perp$, $\mathcal{V}_3 = \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3) \cap \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2)^\perp$, $\mathcal{V}_4 = \mathcal{C}_\Sigma(\mathbf{A}_1 : \mathbf{A}_2 : \mathbf{A}_3)^\perp$, $\mathcal{W}_1 = \mathcal{C}(\mathbf{C}'_3)$, $\mathcal{W}_2 = \mathcal{C}(\mathbf{C}'_2) \cap \mathcal{C}(\mathbf{C}'_3)^\perp$, $\mathcal{W}_3 = \mathcal{C}(\mathbf{C}'_1) \cap \mathcal{C}(\mathbf{C}'_2)^\perp$, $\mathcal{W}_4 = \mathcal{C}(\mathbf{C}'_1)^\perp$.

In practice Σ is not known and should be estimated. A natural way to get an estimator of Σ is to use the sum of squared estimated residuals. If Σ is not structured we estimate the residuals in (4) with

$$\mathbf{R}_r = (\mathbf{I} - \sum_{i=1}^{r-1} \mathbf{P}_{\mathbf{P}_i \mathbf{A}_i, \mathbf{S}_i}) \mathbf{X} (\mathbf{P}_{\mathbf{C}'_{r-1}} - \mathbf{P}_{\mathbf{C}'_r}), \quad r = 1, 2, 3, \dots, m+1,$$

where $\mathbf{P}_i$ and $\mathbf{S}_i$ are given as in Theorem 2.1. Thus a natural estimator of $\mathbf{\Sigma}$ is obtained from the sum of squared residuals, i.e.,

$$n\hat{\mathbf{\Sigma}} = \mathbf{R}_1\mathbf{R}'_1 + \mathbf{R}_2\mathbf{R}'_2 + \cdots + \mathbf{R}_{m+1}\mathbf{R}'_{m+1},$$

which is the maximum likelihood estimator.

3 Estimators when the covariance matrix is linearly structured

In this section we consider the extended growth curve model as in Definition 1.1, but with a linearly structured covariance matrix $\mathbf{\Sigma}$. This $\mathbf{\Sigma}$ will be denoted $\mathbf{\Sigma}^{(s)}$ so that $\mathbf{E} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}^{(s)}, \mathbf{I}_n)$.

The estimation procedure that we propose will rely on the decomposition of the spaces done in Section 2. We will sequentially estimate the inner product in the spaces $\mathcal{V}_i$ ($i = 1, 2, \dots, m$), the residuals in (4), and finally estimate the covariance matrix using all estimated residuals.

To start with, it is natural to use $\mathbf{Q}_1 = \mathbf{H}_1\mathbf{H}'_1$, $\mathbf{H}_1 = \mathbf{X}(\mathbf{I} - \mathbf{P}_{\mathbf{C}'_1})$, to estimate the inner product in the space $\mathcal{V}_1$. We apply the general least squares approach and minimize $\text{tr} \left\{ (\mathbf{Q}_1 - (n - r_1)\mathbf{\Sigma}^{(s)})(\cdot)' \right\}$ with respect to $\mathbf{\Sigma}^{(s)}$, where the notation $(\mathbf{Y})(\cdot)'$ stands for $(\mathbf{Y})(\mathbf{Y})'$. This is done in Lemma 3.1.

By $\text{vec}\mathbf{\Sigma}(K)$ we mean the patterned vectorization of the linearly structured matrix $\mathbf{\Sigma}^{(s)}$, that is the columnwise vectorization of $\mathbf{\Sigma}^{(s)}$ where all 0's and repeated elements (by modulus) have been disregarded. Then there exists a transformation matrix $\mathbf{T}$ such that

$$\text{vec}\mathbf{\Sigma}(K) = \mathbf{T}\text{vec}\mathbf{\Sigma}^{(s)} \text{ or } \text{vec}\mathbf{\Sigma}^{(s)} = \mathbf{T}^+\text{vec}\mathbf{\Sigma}(K), \quad (5)$$

where $\mathbf{T}^+$ denotes the Moore-Penrose generalized inverse of $\mathbf{T}$. Furthermore, from [4], we have

$$\frac{d\mathbf{\Sigma}^{(s)}}{d\mathbf{\Sigma}(K)} = (\mathbf{T}^+)'. \quad (6)$$

Lemma 3.1 *Let $\mathbf{Q}_1 = \mathbf{H}_1\mathbf{H}'_1 = \mathbf{X}(\mathbf{I} - \mathbf{P}_{\mathbf{C}'_1})\mathbf{X}'$, $\hat{\mathbf{\Upsilon}}_1 = (n - r_1)\mathbf{I}$. Then, the minimizer of*

$$f_1(\mathbf{\Sigma}^{(s)}) = \text{tr} \left\{ (\mathbf{Q}_1 - (n - r_1)\mathbf{\Sigma}^{(s)})(\cdot)' \right\}$$

is given by

$$\text{vec}\hat{\mathbf{\Sigma}}_1^{(s)} = \mathbf{T}^+ \left((\mathbf{T}^+)' \hat{\mathbf{\Upsilon}}_1' \hat{\mathbf{\Upsilon}}_1 \mathbf{T}^+ \right)^- (\mathbf{T}^+)' \hat{\mathbf{\Upsilon}}_1' \text{vec}\mathbf{Q}_1.$$

Proof. We may write

$$\begin{aligned} \text{tr} \left\{ \left(\hat{\mathbf{Q}}_1 - (n - r_1)\mathbf{\Sigma}^{(s)} \right) (\cdot)' \right\} &= \left(\text{vec} \left(\hat{\mathbf{Q}}_1 - (n - r_1)\mathbf{\Sigma}^{(s)} \right) \right)' (\cdot) \\ &= \left(\text{vec}\hat{\mathbf{Q}}_1 - \hat{\mathbf{\Upsilon}}_1 \text{vec}\mathbf{\Sigma}^{(s)} \right)' (\cdot), \end{aligned}$$

and thus

$$f_1(\mathbf{\Sigma}^{(s)}) = \left(\text{vec}\hat{\mathbf{Q}}_1 - \hat{\mathbf{\Upsilon}}_1 \text{vec}\mathbf{\Sigma}^{(s)} \right)' (\cdot). \quad (7)$$

Differentiating the expression in the right hand side of (7) with respect to $\text{vec}\Sigma(K)$ and equalizing to $\mathbf{0}$, we get

$$-2\frac{d\Sigma^{(s)}}{d\Sigma(K)}\hat{\mathbf{T}}_1'(\text{vec}\hat{\mathbf{Q}}_1 - \hat{\mathbf{T}}_1\text{vec}\Sigma^{(s)}) = \mathbf{0}. \quad (8)$$

From (5), (6) and (8) we obtain the linear equation

$$(\mathbf{T}^+)' \hat{\mathbf{T}}_1' \hat{\mathbf{T}}_1 \mathbf{T}^+ \text{vec}\Sigma(K) = (\mathbf{T}^+)' \hat{\mathbf{T}}_1' \text{vec}\hat{\mathbf{Q}}_1,$$

which is consistent and a general solution is given by

$$\text{vec}\Sigma(K) = \left((\mathbf{T}^+)' \hat{\mathbf{T}}_1' \hat{\mathbf{T}}_1 \mathbf{T}^+ \right)^- (\mathbf{T}^+)' \hat{\mathbf{T}}_1' \text{vec}\hat{\mathbf{Q}}_1 + ((\mathbf{T}^+)' \hat{\mathbf{T}}_1' \hat{\mathbf{T}}_1 \mathbf{T}^+)^o \mathbf{z},$$

where $\mathbf{z}$ is an arbitrary vector. Hence, using (5) we obtain a unique minimizer of (7) given by

$$\text{vec}\hat{\Sigma}_1^{(s)} = \mathbf{T}^+ \left((\mathbf{T}^+)' \hat{\mathbf{T}}_1' \hat{\mathbf{T}}_1 \mathbf{T}^+ \right)^- (\mathbf{T}^+)' \hat{\mathbf{T}}_1' \text{vec}\hat{\mathbf{Q}}_1, \quad (9)$$

and the proof is complete. $\blacksquare$

Lemma 3.1 gives the first estimator of $\Sigma^{(s)}$. Assuming that $\hat{\Sigma}_1^{(s)}$ is positive definite (which always holds for large n), we can use $\hat{\Sigma}_1^{(s)}$ to define the inner product in the space $\mathcal{V}_1$, and therefore we consider $\mathcal{C}_{\hat{\Sigma}_1^{(s)}}(\mathbf{A}_1)$ instead of $\mathcal{C}_{\Sigma^{(s)}}(\mathbf{A}_1)$. At the same time an estimator of $\mathbf{M}_1$, and also that of $\mathbf{H}_2$ are found by projecting observations on $\mathcal{C}(\mathbf{C}'_1) \otimes \mathcal{V}_1$ and $(\mathcal{C}(\mathbf{C}'_1) \cap \mathcal{C}(\mathbf{C}'_2)^\perp) \otimes \mathcal{V}_1^\perp$ respectively, i.e.,

$$\begin{aligned} \widehat{\mathbf{M}}_1 &= \mathbf{P}_{\mathbf{A}_1, \hat{\Sigma}_1^{(s)}} \mathbf{X} \mathbf{P}_{\mathbf{C}'_1}, \\ \widehat{\mathbf{H}}_2 &= (\mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \hat{\Sigma}_1^{(s)}}) \mathbf{X} (\mathbf{P}_{\mathbf{C}'_1} - \mathbf{P}_{\mathbf{C}'_2}). \end{aligned} \quad (10)$$

A second estimator of $\Sigma^{(s)}$ is obtained using the sum of $\mathbf{Q}_1$ and $\widehat{\mathbf{H}}_2 \widehat{\mathbf{H}}_2'$. Notice that

$$\widehat{\mathbf{H}}_2 \widehat{\mathbf{H}}_2' = (\mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \hat{\Sigma}_1^{(s)}}) \mathbf{X} (\mathbf{P}_{\mathbf{C}'_1} - \mathbf{P}_{\mathbf{C}'_2}) \mathbf{X}' (\mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \hat{\Sigma}_1^{(s)}})'$$

Put $\widehat{\mathbf{T}}_1 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \hat{\Sigma}_1^{(s)}}$ and $\mathbf{W}_1 = \mathbf{X} (\mathbf{P}_{\mathbf{C}'_1} - \mathbf{P}_{\mathbf{C}'_2}) \mathbf{X}'$. Then $\widehat{\mathbf{H}}_2 \widehat{\mathbf{H}}_2' = \widehat{\mathbf{T}}_1 \mathbf{W}_1 \widehat{\mathbf{T}}_1'$ and $\mathbf{Q}_1$ is independent of $\mathbf{W}_1$. Therefore it is natural to condition $\widehat{\mathbf{H}}_2 \widehat{\mathbf{H}}_2'$ with respect to $\mathbf{Q}_1$ and

$$\widehat{\mathbf{H}}_2 \widehat{\mathbf{H}}_2' | \mathbf{Q}_1 \sim W_p \left(\widehat{\mathbf{T}}_1 \Sigma^{(s)} \widehat{\mathbf{T}}_1', r_1 - r_2 \right),$$

where $W_p(\cdot, \cdot)$ stands for the Wishart matrix.

The following lemma gives a second estimator of $\Sigma^{(s)}$ where again the general least squares approach is employed.

Lemma 3.2 Let $\widehat{\mathbf{Q}}_2 = \mathbf{Q}_1 + \widehat{\mathbf{H}}_2 \widehat{\mathbf{H}}_2'$, $\widehat{\mathbf{Y}}_2 = (n - r_1)\mathbf{I} + (r_1 - r_2)\widehat{\mathbf{T}}_1 \otimes \widehat{\mathbf{T}}_1$. Then, the minimizer of

$$f_2(\boldsymbol{\Sigma}^{(s)}) = \text{tr} \left\{ (\widehat{\mathbf{Q}}_2 - [(n - r_1)\boldsymbol{\Sigma}^{(s)} + (r_1 - r_2)\widehat{\mathbf{T}}_1 \boldsymbol{\Sigma}^{(s)} \widehat{\mathbf{T}}_1'])(\cdot)' \right\}$$

is given by

$$\text{vec} \widehat{\boldsymbol{\Sigma}}_2^{(s)} = \mathbf{T}^+ \left((\mathbf{T}^+)' \widehat{\mathbf{Y}}_2' \widehat{\mathbf{Y}}_2 \mathbf{T}^+ \right)^- (\mathbf{T}^+)' \widehat{\mathbf{Y}}_2' \text{vec} \widehat{\mathbf{Q}}_2.$$

Proof. Similar as in Lemma 3.1. ■

Now assume that $\widehat{\boldsymbol{\Sigma}}_2^{(s)}$ is positive definite and use it to define the inner product in $\mathcal{V}_2$, i.e., consider $\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_2^{(s)}}(\widehat{\mathbf{T}}_1 \mathbf{A}_2)$ instead of $\mathcal{C}_{\boldsymbol{\Sigma}^{(s)}}(\mathbf{T}_1 \mathbf{A}_2)$. This gives us an estimator of $\mathbf{M}_2$, and also that of $\mathbf{H}_3$ by projecting observations on $\mathcal{C}(\mathbf{C}'_2) \otimes \mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_2^{(s)}}(\widehat{\mathbf{T}}_1 \mathbf{A}_2)$ and $(\mathcal{C}(\mathbf{C}'_2) \cap \mathcal{C}(\mathbf{C}'_3)^\perp) \otimes (\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_1^{(s)}}(\mathbf{A}_1) + \mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_2^{(s)}}(\widehat{\mathbf{T}}_1 \mathbf{A}_2))^\perp$ respectively, i.e.,

$$\begin{aligned} \widehat{\mathbf{M}}_2 &= \mathbf{P}_{\widehat{\mathbf{T}}_1 \mathbf{A}_2, \widehat{\boldsymbol{\Sigma}}_2^{(s)}} \mathbf{X} \mathbf{P}_{\mathbf{C}'_2}, \\ \widehat{\mathbf{H}}_3 &= \widehat{\mathbf{T}}_2 \mathbf{X} (\mathbf{P}_{\mathbf{C}'_2} - \mathbf{P}_{\mathbf{C}'_3}), \\ \widehat{\mathbf{T}}_2 &= \mathbf{I} - \mathbf{P}_{\mathbf{A}_1, \widehat{\boldsymbol{\Sigma}}_1^{(s)}} - \mathbf{P}_{\widehat{\mathbf{T}}_1 \mathbf{A}_2, \widehat{\boldsymbol{\Sigma}}_2^{(s)}}. \end{aligned}$$

To derive a third estimator of $\boldsymbol{\Sigma}^{(s)}$, the idea is to use the sum $\widehat{\mathbf{Q}}_3 = \widehat{\mathbf{Q}}_2 + \widehat{\mathbf{H}}_3 \widehat{\mathbf{H}}_3'$ in a similar way as in Lemma 3.1 or Lemma 3.2. We continue the same process until all residuals are estimated and then use the sum of squared estimated residuals to estimate the covariance matrix that can be understood as a dispersion matrix. After the $(r - 1)^{\text{th}}$ stage of the process we have already $r - 1$ estimates of $\boldsymbol{\Sigma}^{(s)}$. In Lemma 3.3, we show how to obtain the r^{th} estimate of $\boldsymbol{\Sigma}^{(s)}$. Before we state it, we notice that after the $(r - 1)^{\text{th}}$ stage we have also the following quantities:

$$\begin{aligned} \mathbf{W}_i &= \mathbf{X} (\mathbf{P}_{\mathbf{C}'_i} - \mathbf{P}_{\mathbf{C}'_{i+1}}) \mathbf{X}' \sim W_p(\boldsymbol{\Sigma}^{(s)}, r_i - r_{i+1}), \\ i &= 0, 1, 2, \dots, m, \end{aligned} \tag{11}$$

$$\widehat{\mathbf{P}}_j = \widehat{\mathbf{T}}_{j-1} \widehat{\mathbf{T}}_{j-2} \times \dots \times \widehat{\mathbf{T}}_0, \widehat{\mathbf{T}}_0 = \mathbf{I}, j = 1, 2, \dots, r - 1, \tag{12}$$

$$\widehat{\mathbf{T}}_i = \mathbf{I} - \mathbf{P}_{\widehat{\mathbf{P}}_i \mathbf{A}_i, \widehat{\boldsymbol{\Sigma}}_i^{(s)}}, i = 1, 2, \dots, r - 1, \tag{13}$$

$$\widehat{\mathbf{T}}_i = \mathbf{I} - \sum_{j=1}^i \mathbf{P}_{\widehat{\mathbf{P}}_j \mathbf{A}_j, \widehat{\boldsymbol{\Sigma}}_j^{(s)}}, i = 1, 2, \dots, r - 1, \tag{14}$$

$$\widehat{\mathbf{H}}_i \widehat{\mathbf{H}}_i' = \widehat{\mathbf{T}}_{i-1} \mathbf{W}_{i-1} \widehat{\mathbf{T}}_{i-1}', \widehat{\mathbf{T}}_0 = \mathbf{I}, i = 1, 2, \dots, r, \tag{15}$$

$$\begin{aligned} \widehat{\mathbf{Q}}_i &= \sum_{j=1}^i \widehat{\mathbf{H}}_j \widehat{\mathbf{H}}_j' = \sum_{j=1}^i \widehat{\mathbf{T}}_{j-1} \mathbf{W}_{j-1} \widehat{\mathbf{T}}_{j-1}', \\ i &= 1, 2, \dots, r, \end{aligned} \tag{16}$$

$$\widehat{\mathbf{Y}}_i = \sum_{j=1}^i (r_{j-1} - r_j) \widehat{\mathbf{T}}_{j-1} \otimes \widehat{\mathbf{T}}_{j-1}, i = 1, 2, \dots, r. \tag{17}$$

Lemma 3.3 Let $\widehat{\mathbf{T}}_i$, $\widehat{\mathbf{Q}}_r$ and $\widehat{\mathbf{Y}}_r$ be defined as in (14), (16) and (17), respectively. Then, the minimizers of

$$f_r(\boldsymbol{\Sigma}^{(s)}) = \text{tr} \left\{ \left(\widehat{\mathbf{Q}}_r - \sum_{i=1}^r (r_{i-1} - r_i) \widehat{\mathbf{T}}_{i-1} \boldsymbol{\Sigma}^{(s)} \widehat{\mathbf{T}}_{i-1}' \right) ()' \right\}, \quad r = 1, 2, \dots, m+1,$$

are given by

$$\text{vec} \widehat{\boldsymbol{\Sigma}}_r^{(s)} = \mathbf{T}^+ \left((\mathbf{T}^+)' \widehat{\mathbf{Y}}_r' \widehat{\mathbf{Y}}_r \mathbf{T}^+ \right)^- (\mathbf{T}^+)' \widehat{\mathbf{Y}}_r' \text{vec} \widehat{\mathbf{Q}}_r.$$

Proof. Similar as in Lemma 3.1. ■

Lemma 3.3 gives the r^{th} estimate of $\boldsymbol{\Sigma}^{(s)}$ and the $(m+1)^{\text{th}}$ estimate of $\boldsymbol{\Sigma}^{(s)}$ can be understood as a dispersion matrix as it uses all information that we have in residuals. Now the results at our disposal permit us to propose estimators of parameters in the extended growth curve model with a linearly structured covariance matrix.

Theorem 3.4 Let the extended growth curve model be given by (1.1). Then

(i) A consistent estimator of the structured covariance matrix $\boldsymbol{\Sigma}^{(s)}$ is given by

$$\text{vec} \widehat{\boldsymbol{\Sigma}}_{m+1}^{(s)} = \mathbf{T}^+ \left((\mathbf{T}^+)' \widehat{\mathbf{Y}}_{m+1}' \widehat{\mathbf{Y}}_{m+1} \mathbf{T}^+ \right)^- (\mathbf{T}^+)' \widehat{\mathbf{Y}}_{m+1}' \text{vec} \widehat{\mathbf{Q}}_{m+1}. \quad (18)$$

(ii) An unbiased estimator of the mean is given by

$$\widehat{E[\mathbf{X}]} = \sum_{i=1}^m (\mathbf{I} - \widehat{\mathbf{T}}_i) \mathbf{X} \mathbf{C}_i' (\mathbf{C}_i \mathbf{C}_i')^- \mathbf{C}_i. \quad (19)$$

where $\widehat{\mathbf{T}}_i$, $\widehat{\mathbf{Q}}_r$ and $\widehat{\mathbf{Y}}_r$ be defined as in (13), (16) and (17), respectively.

Proof. (i) The consistency of the estimator in (18) is established through the following implications (the notation ‘ $\xrightarrow{p}$ ’ means convergence in probability)

$$\begin{array}{ccc} \frac{1}{n-r_1} \mathbf{Q}_1 & \xrightarrow{p} & \boldsymbol{\Sigma}^{(s)} \\ \Downarrow & & \\ \widehat{\boldsymbol{\Sigma}}_1^{(s)} & \xrightarrow{p} & \boldsymbol{\Sigma}^{(s)} \\ \Downarrow & & \\ \widehat{\boldsymbol{\Sigma}}_2^{(s)} & \xrightarrow{p} & \boldsymbol{\Sigma}^{(s)} \\ \Downarrow & & \\ \vdots & & \\ \Downarrow & & \\ \widehat{\boldsymbol{\Sigma}}_{m+1}^{(s)} & \xrightarrow{p} & \boldsymbol{\Sigma}^{(s)}. \end{array}$$

The above implications can be easily checked. The first line follows from the well known fact that the (corrected) sample covariance matrix is a consistent estimator of the true covariance matrix. The rest is established using Cramer-Slutsky's theorem [14]. (ii) Next we establish the unbiasedness of the estimator given by (19). From Theorem 2.2. (i) and using uniqueness property of projectors it is possible to rewrite the estimated mean as

$$\widehat{E[\mathbf{X}]} = \sum_{i=1}^m \mathbf{P}_{\mathcal{A}_i, \widehat{\Sigma}_i^{(s)}} \mathbf{X} (\mathbf{P}_{C'_i} - \mathbf{P}_{C'_{i+1}}).$$

Using the linearity of the expectation operator and independence of $\mathbf{P}_{\mathcal{A}_i, \widehat{\Sigma}_i^{(s)}}$ and $\mathbf{X}(\mathbf{P}_{C'_i} - \mathbf{P}_{C'_{i+1}})$, we have

$$\begin{aligned} E[\widehat{E[\mathbf{X}]}] &= \sum_{i=1}^m E \left[\mathbf{P}_{\mathcal{A}_i, \widehat{\Sigma}_i^{(s)}} \right] E \left[\mathbf{X}(\mathbf{P}_{C'_i} - \mathbf{P}_{C'_{i+1}}) \right] \\ &= E \left[\mathbf{P}_{\mathcal{A}_m, \widehat{\Sigma}_m^{(s)}} \right] E \left[\mathbf{X} \mathbf{P}_{C'_m} \right] \\ &\quad + E \left[\mathbf{P}_{\mathcal{A}_{m-1}, \widehat{\Sigma}_{m-1}^{(s)}} \right] E \left[\mathbf{X}(\mathbf{P}_{C'_{m-1}} - \mathbf{P}_{C'_m}) \right] \\ &\quad + E \left[\mathbf{P}_{\mathcal{A}_{m-2}, \widehat{\Sigma}_{m-2}^{(s)}} \right] E \left[\mathbf{X}(\mathbf{P}_{C'_{m-2}} - \mathbf{P}_{C'_{m-1}}) \right] + \cdots \\ &\quad + E \left[\mathbf{P}_{\mathcal{A}_1, \widehat{\Sigma}_1^{(s)}} \right] E \left[\mathbf{X}(\mathbf{P}_{C'_1} - \mathbf{P}_{C'_2}) \right]. \end{aligned}$$

Since $E[\mathbf{X}] = \sum_{i=1}^m \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i$, using the facts that $\mathcal{C}(C'_i) \subseteq \mathcal{C}(C'_{i-1})$ and

$$\mathbf{P}_{\mathcal{A}_n, \widehat{\Sigma}_n^{(s)}} \left(\sum_{i=1}^n \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) = \sum_{i=1}^n \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i,$$

we get

$$\begin{aligned} E[\widehat{E[\mathbf{X}]}] &= \left(\sum_{i=1}^m \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) \mathbf{P}_{C'_m} + \left(\sum_{i=1}^{m-1} \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) (\mathbf{P}_{C'_{m-1}} - \mathbf{P}_{C'_m}) + \cdots \\ &\quad + \left(\sum_{i=1}^2 \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) (\mathbf{P}_{C'_2} - \mathbf{P}_{C'_3}) + \mathbf{A}_1 \mathbf{B}_1 \mathbf{C}_1 (\mathbf{P}_{C'_1} - \mathbf{P}_{C'_2}) \\ &= \mathbf{A}_m \mathbf{B}_m \mathbf{C}_m + \left(\sum_{i=1}^{m-1} \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) \mathbf{P}_{C'_m} + \mathbf{A}_{m-1} \mathbf{B}_{m-1} \mathbf{C}_{m-1} \\ &\quad + \left(\sum_{i=1}^{m-2} \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) \mathbf{P}_{C'_{m-1}} - \left(\sum_{i=1}^{m-1} \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) \mathbf{P}_{C'_m} + \cdots \\ &\quad + \mathbf{A}_2 \mathbf{B}_2 \mathbf{C}_2 + \mathbf{A}_1 \mathbf{B}_1 \mathbf{C}_1 \mathbf{P}_{C'_2} - \left(\sum_{i=1}^2 \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i \right) \mathbf{P}_{C'_3} \\ &\quad + \mathbf{A}_1 \mathbf{B}_1 \mathbf{C}_1 - \mathbf{A}_1 \mathbf{B}_1 \mathbf{C}_1 \mathbf{P}_{C'_2}. \end{aligned}$$

Canceling the opposite terms in the last expression, we get the desired result. ■

4 Simulation study

The aim of this section is to investigate the performance of the proposed algorithm for small sample sizes. In fact, for some structures of the covariance matrix, the proposed technique may produce non positive definite estimates for small sample sizes. In our simulations we will also study the same problem with respect to the number of profiles m .

4.1 Procedures and methods

We consider the following three scenarios:

Scenario S1: The covariance matrix has the uniform structure:

$$\boldsymbol{\Sigma}^{(s)} = 3[(1 - 0.7)\mathbf{I}_4 + 0.7\mathbf{J}_4],$$

where $\mathbf{I}_4$ and $\mathbf{J}_4$ stand for the 4×4 identity matrix and the 4×4 matrix of ones respectively.

Scenario S2: The covariance matrix has the Toeplitz structure with different variances:

$$\boldsymbol{\Sigma}^{(s)} = \begin{pmatrix} 4 & 1.2 & 0.8 & 0.9 \\ 1.2 & 5 & 1.2 & 0.8 \\ 0.8 & 1.2 & 6 & 1.2 \\ 0.9 & 0.8 & 1.2 & 7 \end{pmatrix}.$$

Scenario S3: The covariance matrix has the banded structure:

$$\boldsymbol{\Sigma}^{(s)} = \begin{pmatrix} 8 & 2 & 0 & 0 \\ 2 & 10 & 4 & 0 \\ 0 & 4 & 12 & 6 \\ 0 & 0 & 6 & 14 \end{pmatrix}.$$

In all scenarios the number of simulation runs was set to 5000 for each sample size considered ($n = 12, 24, 36, 60, 72, 240, 360$), and the percentage of non positive definite estimates of the covariance matrix was calculated. Data is generated from $\mathbf{X} \sim N_{p,n}(\sum_{i=1}^m \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i, \boldsymbol{\Sigma}, \mathbf{I})$, $m = 1, 2, 3$, with $p = 4$. The within design matrices are

$$\begin{aligned} \mathbf{A}'_1 &= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \mathbf{A}'_2 = \begin{pmatrix} 1^2 & 2^2 & 3^2 & 4^2 \end{pmatrix}, \\ \mathbf{A}'_3 &= \begin{pmatrix} 1^3 & 2^3 & 3^3 & 4^3 \end{pmatrix}. \end{aligned}$$

The between design matrices are chosen so that we have the same number of individuals in each group (two group for $m = 1, 2$ and three groups for $m = 3$). For example, for $m = 3$, we have

$$\begin{aligned} \mathbf{C}_1 &= \left(\mathbf{1}'_{n/3} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} : \mathbf{1}'_{n/3} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : \mathbf{1}'_{n/3} \otimes \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right), \\ \mathbf{C}_2 &= \left(\mathbf{1}'_{n/3} \otimes \begin{pmatrix} 0 \\ 0 \end{pmatrix} : \mathbf{1}'_{n/3} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \mathbf{1}'_{n/3} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right), \\ \mathbf{C}_3 &= \left(\mathbf{0}'_{n/3} : \mathbf{0}'_{n/3} : \mathbf{1}'_{n/3} \right), \end{aligned}$$

Table 1: Percentage of non positive definite estimates of different linearly structured covariance matrices according to the number of profiles (m) and the sample size (n).

Number of profiles m	Sample size n	Linear structure		
		S1	S2	S3
1	12	0	5.48	7.82
	24	0	0.04	0.64
	36	0	0	0.26
	60	0	0	0
	72	0	0	0
	240	0	0	0
	360	0	0	0
2	12	0	12.56	15.12
	24	0	0.38	1.54
	36	0	0.02	0.24
	60	0	0	0.02
	72	0	0	0
	240	0	0	0
	360	0	0	0
3	12	0	86.98	63.84
	24	0	39.64	14.84
	36	0	12.98	4.08
	60	0	1.60	0.22
	72	0	0.50	0.10
	240	0	0	0
	360	0	0	0

where the notation $\mathbf{1}_k$ (respectively $\mathbf{0}_k$) stands for a column vector of 1's (respectively of 0's) of size k . The parameter matrices are

$$\mathbf{B}_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} 2 & 3 \end{pmatrix}, \quad \mathbf{B}_3 = 4.$$

4.2 Results

Table 1 contains percentages of non positive definite estimates of the covariance matrix for each scenario considered. One can see that for the uniform structure our techniques perform well as we have 100% of positive definite estimates in all cases. Also we have good performance for the other two cases where we have small percentages of non definite estimates for $n < 36$ whether we have one profile or two profiles. However, for three profiles case we have high percentage of non positive definite estimates for $n < 36$. Hence, concerning the number of profiles m , we note that as m increases the percentages of non positive definite estimates also increase except for the case of uniform structure.

References

- [1] R. F. Pothoff and S. N. Roy, “A generalized multivariate analysis of variance model useful especially for growth curve problems,” *Biometrika*, vol. 51, pp. 313–326, 1964.
- [2] A. P. Verbyla and W. N. Venables, “An extension of the growth curve model,” *Biometrika*, vol. 75, no. 1, pp. 129–138, 1988.
- [3] D. von Rosen, “Maximum likelihood estimators in multivariate linear normal models,” *Journal of multivariate analysis*, vol. 31, no. 2, pp. 187–200, 1989.
- [4] T. Kollo and D. von Rosen, *Advanced Multivariate Statistics with Matrices*. Dordrecht, The Netherlands: Springer, 2005.
- [5] K. Filipiak and D. von Rosen, “On MLEs in an extended multivariate linear growth curve model,” *Metrika*, vol. 75, no. 8, pp. 1069–1092, 2012.
- [6] J. Hu, “Properties of the explicit estimators in the extended growth curve model,” *Statistics*, vol. 44, no. 5, pp. 477–492, 2010.
- [7] J. Hu, G. Yan, and J. You, “Estimation for an additive growth curve model with orthogonal design matrices,” *Bernoulli*, vol. 17, no. 4, pp. 1400–1419, 2011.
- [8] J. C.-S. Lee and S. Geisser, “Applications of growth curve prediction,” *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 37, no. 2, pp. 239–256, 1975.
- [9] J. C. Lee, “Prediction and estimation of growth curves with special covariance structures,” *Journal of the American Statistical Association*, vol. 83, no. 402, pp. 432–440, 1988.
- [10] C. G. Khatri, “Testing some covariance structures under a growth curve model,” *Journal of Multivariate Analysis*, vol. 3, no. 1, pp. 102–116, 1973.
- [11] M. Ohlson and D. von Rosen, “Explicit estimators of parameters in the growth curve model with linearly structured covariance matrices,” *Journal of Multivariate Analysis*, vol. 101, no. 5, pp. 1284–1295, 2010.
- [12] J. Nzabanita, M. Singull, and D. von Rosen, “Estimation of parameters in the extended growth curve model with a linearly structured covariance matrix,” *Acta et Commentationes Universitatis Tartuensis de Mathematica*, vol. 16, no. 1, pp. 13–32, 2012.
- [13] C. G. Khatri, “Some results for the singular normal multivariate regression models,” *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 267–280, 1968.
- [14] H. Cramér, *Mathematical Methods in Statistics*. Princeton: Princeton University Press, 1946.