ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 16, Number 1, 2012
Available online at www.math.ut.ee/acta/

Estimation of parameters in the extended growth curve model with a linearly structured covariance matrix

Joseph Nzabanita, Dietrich von Rosen, and Martin Singull

Abstract. In this paper the extended growth curve model with two terms and a linearly structured covariance matrix is considered. We propose an estimation procedure that handles linearly structured covariance matrices. The idea is first to estimate the covariance matrix when finding the inner product in a regression space and thereafter re-estimate it when it should be interpreted as a dispersion matrix. This idea is exploited by decomposing the residual space, the orthogonal complement to the design space, into three orthogonal subspaces. Studying residuals obtained from projections of observations on these subspaces yields explicit consistent estimators of the covariance matrix. An explicit consistent estimator of the mean is also proposed and numerical examples are given.

1. Introduction

Growth curve analysis is a topic with many applications in different fields such as medicine, natural sciences, social sciences, etc. The growth curve model was introduced in [7]. Since then many authors have been interested by the model and renown follow-up papers are [8] and [2]. In this paper we study the extended growth curve model with two terms and a linearly structured covariance matrix. The extended growth curve model was introduced in [9] and may be defined as follows:

Definition 1.1 (Extended growth curve model)**.** Let $\mathbf{X} : p\times n$, $\mathbf{A}_i : p\times q_i$, $\mathbf{B}_i : q_i \times k_i$, $\mathbf{C}_i : k_i \times n$, $r(\mathbf{C}_1) + p \leq n$, $i = 1, 2, \ldots, m$, $\mathcal{C}(\mathbf{C}'_i) \subseteq \mathcal{C}(\mathbf{C}'_{i-1})$, $i = 2, 3, \ldots, m$, where $r(\cdot)$ and $\mathcal{C}(\cdot)$ represent the rank and column space of

---

Received October 24, 2011.
2010 *Mathematics Subject Classification.* 62H12.
Key words and phrases. Extended growth curve model, estimation, linearly structured covariance matrix, residuals.

a matrix, respectively. 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 main object of this paper is to derive explicit estimators of parameters in the extended growth curve model with two terms and a linearly structured covariance matrix, which means that, see [3], for $\mathbf{\Sigma} = (\sigma_{ij})$ 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 $|\sigma_{ij}| = |\sigma_{kl}| \neq 0$.

Linear structures for the covariance matrix exist very often in statistical applications. These structures are, for example, the uniform structure (or intraclass structure) considered for the first time in [13], the compound symmetry structure which is an extension of the uniform structure [12]. Other linear structures for the covariance matrix often encountered are the matrix with zeros, for example a banded covariance matrix [4], the Toeplitz or circular Toeplitz [6], etc.

In general there is no problem to estimate the covariance matrix when it is completely unknown. However, problems arise when one has to take into account that there exists a structure generated by a few number of parameters. For the unstructured case, several approaches to find estimators of parameters in the growth curve model or extended growth curve model exist. One of those approaches is the maximum likelihood method. The maximum likelihood estimators of parameters in the growth curve model have been studied by many authors, see for instance [11] and [9]. For the extended growth curve model as given in Definition 1.1 an exhaustive description of how to get those estimators can be found in [3]. For the structured case we can use, in principle, a maximum likelihood approach. However, this will no longer give any explicit estimator even for the intraclass structure, which is among the simplest cases. We must therefore rely on iterative methods. When data sets are very large, iterative methods perform poorly and non-iterative methods become of great interest [5]. In [5], the classical growth curve model with a linearly structured covariance matrix was studied and explicit estimators were proposed. Our aim here is to obtain explicit estimators in the extended growth curve model with two terms and a linearly structured covariance matrix. We propose an estimation procedure that handles linear structured covariance matrices. That procedure is somewhat based on the maximum likelihood method for the unstructured case and Section 2 is devoted to that.

2. Maximum likelihood estimators

In this section we consider the extended growth curve model as given in Definition 1.1 when $m = 2$. In this case, under general settings, the maximum likelihood estimators for the parameter matrices $\mathbf{B}_1$ and $\mathbf{B}_2$, see e.g. [3] for derivation, are given by

$$\widehat{\mathbf{B}}_2 = (\mathbf{A}_2'\mathbf{P}_2'\mathbf{S}_2^{-1}\mathbf{P}_2\mathbf{A}_2)^-\mathbf{A}_2'\mathbf{P}_2'\mathbf{S}_2^{-1}\mathbf{X}\mathbf{C}_2'(\mathbf{C}_2\mathbf{C}_2')^- + (\mathbf{A}_2'\mathbf{P}_2)^o\mathbf{Z}_{21} + \mathbf{A}_2'\mathbf{Z}_{22}\mathbf{C}_2^{o\prime},$$
$$\widehat{\mathbf{B}}_1 = (\mathbf{A}_1'\mathbf{S}_1^{-1}\mathbf{A}_1)^-\mathbf{A}_1'\mathbf{S}_1^{-1}(\mathbf{X} - \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2)\mathbf{C}_1'(\mathbf{C}_1\mathbf{C}_1')^- + \mathbf{A}_1^{o\prime}\mathbf{Z}_{11} + \mathbf{A}_1'\mathbf{Z}_{12}\mathbf{C}_1^{o\prime},$$

where

$$\begin{aligned}
\mathbf{S}_1 &= \mathbf{X}\left(\mathbf{I} - \mathbf{C}_1'(\mathbf{C}_1\mathbf{C}_1')^-\mathbf{C}_1\right)\mathbf{X}', \\
\mathbf{P}_2 &= \mathbf{I} - \mathbf{A}_1(\mathbf{A}_1'\mathbf{S}_1^{-1}\mathbf{A}_1)^-\mathbf{A}_1'\mathbf{S}_1^{-1}, \\
\mathbf{S}_2 &= \mathbf{S}_1 + \mathbf{P}_2\mathbf{X}\mathbf{C}_1'(\mathbf{C}_1\mathbf{C}_1')^-\mathbf{C}_1\left(\mathbf{I} - \mathbf{C}_2'(\mathbf{C}_2\mathbf{C}_2')^-\mathbf{C}_2\right)\mathbf{C}_1'(\mathbf{C}_1\mathbf{C}_1')^-\mathbf{C}_1\mathbf{X}'\mathbf{P}_2',
\end{aligned}$$

$\mathbf{Z}_{kl}$ are arbitrary matrices, $\mathbf{C}^o$ stands for any matrix of full rank spanning $\mathcal{C}(\mathbf{C})^\perp$, $^\perp$ denotes the orthogonal complement, and $\mathbf{G}^-$ denotes an arbitrary generalized inverse in the sense that $\mathbf{G}\mathbf{G}^-\mathbf{G} = \mathbf{G}$.

Assuming that matrices $\mathbf{A}_i$, $\mathbf{C}_i$ are of full rank and that $\mathcal{C}(\mathbf{A}_1) \cap \mathcal{C}(\mathbf{A}_2) = \{0\}$, the unique maximum likelihood estimators are

$$\begin{aligned}
\widehat{\mathbf{B}}_2 &= (\mathbf{A}_2'\mathbf{P}_2'\mathbf{S}_2^{-1}\mathbf{P}_2\mathbf{A}_2)^{-1}\mathbf{A}_2'\mathbf{P}_2'\mathbf{S}_2^{-1}\mathbf{X}\mathbf{C}_2'(\mathbf{C}_2\mathbf{C}_2')^{-1}, \\
\widehat{\mathbf{B}}_1 &= (\mathbf{A}_1'\mathbf{S}_1^{-1}\mathbf{A}_1)^{-1}\mathbf{A}_1'\mathbf{S}_1^{-1}(\mathbf{X} - \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2)\mathbf{C}_1'(\mathbf{C}_1\mathbf{C}_1')^{-1}.
\end{aligned}$$

Obviously, under general settings, the maximum likelihood estimators $\widehat{\mathbf{B}}_1$ and $\widehat{\mathbf{B}}_2$ are not unique due to the arbitrariness of matrices $\mathbf{Z}_{kl}$. However, it is worth noting that the estimated mean

$$\widehat{E[\mathbf{X}]} = \mathbf{A}_1\widehat{\mathbf{B}}_1\mathbf{C}_1 + \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2$$

is always unique and therefore $\widehat{\mathbf{\Sigma}}$ given by

$$n\widehat{\mathbf{\Sigma}} = (\mathbf{X} - \mathbf{A}_1\widehat{\mathbf{B}}_1\mathbf{C}_1 - \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2)(\mathbf{X} - \mathbf{A}_1\widehat{\mathbf{B}}_1\mathbf{C}_1 - \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2)' \qquad (1)$$

is also unique.

Now to shorten matrix expressions we introduce different notations that we will use throughout this paper. For any pair of matrices $\mathbf{S}$ and $\mathbf{A}$, where $\mathbf{S}$ is positive definite, we define

$$\mathbf{P}_{\mathbf{A},\mathbf{S}} = \mathbf{A}(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^-\mathbf{A}'\mathbf{S}^{-1}.$$

It can be shown, see [3], that $\mathbf{P}_{\mathbf{A},\mathbf{S}}$ is a projector and that

$$\mathbf{P}_{\mathbf{A},\mathbf{S}} = \mathbf{I} - \mathbf{P}'_{\mathbf{A}^o,\mathbf{S}^{-1}} = \mathbf{I} - \mathbf{S}\mathbf{A}^o(\mathbf{A}^{o\prime}\mathbf{S}\mathbf{A}^o)^-\mathbf{A}^{o\prime}.$$

Moreover, $\mathbf{P}^o_{\mathbf{A},\mathbf{S}} = \mathbf{I} - \mathbf{P}'_{\mathbf{A},\mathbf{S}} = \mathbf{P}_{\mathbf{A}^o,\mathbf{S}^{-1}}$. If $\mathbf{S} = \mathbf{I}$, instead of $\mathbf{P}_{\mathbf{A},\mathbf{I}}$ the notation $\mathbf{P}_{\mathbf{A}}$ will be used.

With these notations, it is straightforward to check that the estimated mean may be written as

$$\mathbf{A}_1\widehat{\mathbf{B}}_1\mathbf{C}_1 + \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2 = \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1}\mathbf{X}\mathbf{P}_{\mathbf{C}'_1} + \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{S}_2}\mathbf{X}\mathbf{P}_{\mathbf{C}'_2}, \tag{2}$$

where $\mathbf{T}_1 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1}$.

This structure reveals a connection to a paradigm used by least squares or maximum likelihood estimation procedures connected to linear models. These procedures use the following paradigm: estimators of the mean parameters are based on a projection of the observations on the space generated by the design matrix whereas estimators of the variance parameters are based on a projection of the observations on the orthogonal complement to the design space. In principal this also takes place for the extended growth curve model, although there are some complications.

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

$$\mathbf{A}_1\widehat{\mathbf{B}}_1\mathbf{C}_1 + \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2 = \mathbf{P}_{\mathbf{A}_1,\mathbf{\Sigma}}\mathbf{X}\mathbf{P}_{\mathbf{C}'_1} + \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{\Sigma}}\mathbf{X}\mathbf{P}_{\mathbf{C}'_2},$$

where $\mathbf{T}_1 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\mathbf{\Sigma}}$. Thus, we see that in the projections, if $\mathbf{\Sigma}$ is unknown, the parameter has been replaced with either $\mathbf{S}_1$ or $\mathbf{S}_2$, which according to their expressions are not maximum likelihood estimators. However, it can be shown that $n^{-1}\mathbf{S}_1 \to \mathbf{\Sigma}$ in probability and $n^{-1}\mathbf{S}_2 \to \mathbf{\Sigma}$ in probability.

To summarize we see that estimation is performed through projections on certain subspaces and indeed the decomposition is essential for both the model (parameter) interpretation and the model evaluation. Now we indicate how we make such a decomposition that suits our purposes. Applying the vec-operator on both sides of (2) we get

$$\begin{aligned}(\mathbf{C}'_1 \otimes \mathbf{A}_1)\mathrm{vec}\widehat{\mathbf{B}}_1 \quad &+ \quad (\mathbf{C}'_2 \otimes \mathbf{A}_2)\mathrm{vec}\widehat{\mathbf{B}}_2 \\ &= \quad [(\mathbf{P}_{\mathbf{C}'_1} \otimes \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1}) + (\mathbf{P}_{\mathbf{C}'_2} \otimes \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{S}_2})]\mathrm{vec}\mathbf{X},\end{aligned}$$

where $\otimes$ denotes the Kronecker product.

Note that the matrix $\mathbf{P} = (\mathbf{P}_{\mathbf{C}'_1} \otimes \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1}) + (\mathbf{P}_{\mathbf{C}'_2} \otimes \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{S}_2})$ is a projector. Moreover,

$$\mathcal{C}(\mathbf{P}) = \mathcal{C}(\mathbf{C}'_1) \otimes \mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1) + \mathcal{C}(\mathbf{C}'_2) \otimes \mathcal{C}_{\mathbf{S}_2}(\mathbf{T}_1\mathbf{A}_2), \tag{3}$$

where now $\otimes$ denotes a tensor product of linear spaces. Therefore $\mathcal{C}(\mathbf{P})$ is used to estimate $\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2$ whereas $\mathcal{C}(\mathbf{P})^\perp$ is used to create residuals. Figure 1 illustrates these spaces; it shows that three regions are describing the mean and six others are describing the residuals.

Notice that the sum in (3) is actually an orthogonal direct sum of two subspaces. Therefore we may consider the following decomposition of the

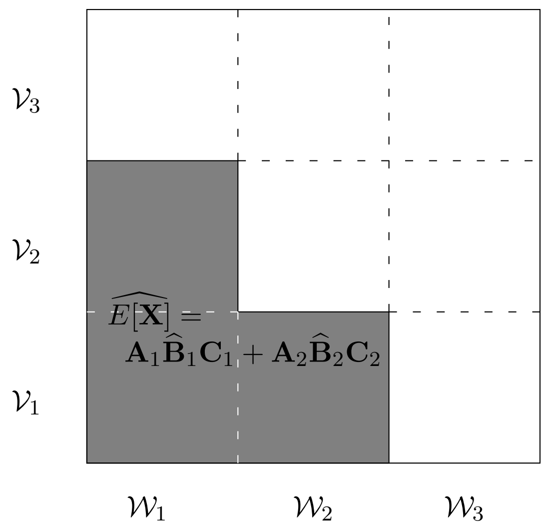

FIGURE 1. Decomposition of the whole space according to the within and between individuals designs illustrating the mean and residual spaces: $\mathcal{V}_1 = \mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1)$, $\mathcal{V}_2 = \mathcal{C}_{\mathbf{S}_2}(\mathbf{T}_1\mathbf{A}_2)$, $\mathcal{V}_3 = (\mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1)+\mathcal{C}_{\mathbf{S}_2}(\mathbf{T}_1\mathbf{A}_2))^{\perp}$, $\mathcal{W}_1 = (\mathcal{C}(\mathbf{C}_2')$, $\mathcal{W}_2 = (\mathcal{C}(\mathbf{C}_1')\cap \mathcal{C}(\mathbf{C}_2')^{\perp}$, $\mathcal{W}_3 = (\mathcal{C}(\mathbf{C}_1')^{\perp}$.

estimated mean

$$\mathbf{A}_1\widehat{\mathbf{B}}_1\mathbf{C}_1 + \mathbf{A}_2\widehat{\mathbf{B}}_2\mathbf{C}_2 = \mathbf{M}_1 + \mathbf{M}_2,$$

where

$$\mathbf{M}_1 = \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1}\mathbf{X}\mathbf{P}_{\mathbf{C}_1'} \text{ and } \mathbf{M}_2 = \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{S}_2}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}.$$

To estimate $\mathbf{\Sigma}$, the general idea is to use the variation in the residuals. For our purposes we decompose the residual space into three orthogonal subspaces, see Figure 2, as

$$\mathcal{C}(\mathbf{P})^{\perp} = \mathbf{I} \boxplus \mathbf{II} \boxplus \mathbf{III},$$

where

$$\begin{aligned}
\mathbf{I} &= \mathcal{C}(\mathbf{C}_1')^{\perp} \otimes \mathcal{V}, \\
\mathbf{II} &= \left(\mathcal{C}(\mathbf{C}_1') \cap \mathcal{C}(\mathbf{C}_2')^{\perp}\right) \otimes \mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1)^{\perp}, \\
\mathbf{III} &= \mathcal{C}(\mathbf{C}_2') \otimes (\mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1) + \mathcal{C}_{\mathbf{S}_2}(\mathbf{T}_1\mathbf{A}_2))^{\perp},
\end{aligned}$$

$\mathcal{V}$ represents the whole space and $\boxplus$ denotes the orthogonal direct sum of tensor spaces.

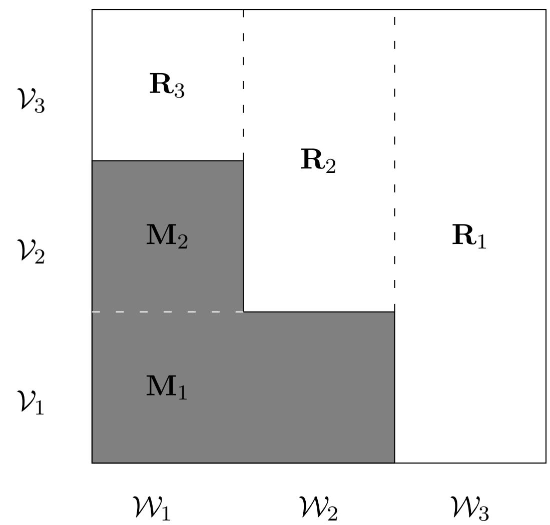

FIGURE 2. Decomposition of the whole space according to the within and between individuals designs illustrating the mean and residual spaces: $\mathcal{V}_1 = \mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1)$, $\mathcal{V}_2 = \mathcal{C}_{\mathbf{S}_2}(\mathbf{T}_1\mathbf{A}_2)$, $\mathcal{V}_3 = (\mathcal{C}_{\mathbf{S}_1}(\mathbf{A}_1) + \mathcal{C}_{\mathbf{S}_2}(\mathbf{T}_1\mathbf{A}_2))^{\perp}$, $\mathcal{W}_1 = (\mathcal{C}(\mathbf{C}_2')$, $\mathcal{W}_2 = (\mathcal{C}(\mathbf{C}_1') \cap \mathcal{C}(\mathbf{C}_2')^{\perp}$, $\mathcal{W}_3 = (\mathcal{C}(\mathbf{C}_1')^{\perp}$.

The residuals obtained by projecting data to these subspaces are respectively

$$\begin{aligned}
\mathbf{R}_1 &= \mathbf{X}(\mathbf{I} - \mathbf{P}_{\mathbf{C}_1'}),\\
\mathbf{R}_2 &= (\mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1})\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'} - \mathbf{P}_{\mathbf{C}_2'}),\\
\mathbf{R}_3 &= (\mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\mathbf{S}_1} - \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{S}_2})\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}.
\end{aligned}$$

Thus a natural estimator of $\mathbf{\Sigma}$ is obtained from the sum of squared residuals, i.e.,

$$n\widehat{\mathbf{\Sigma}} = \mathbf{R}_1\mathbf{R}_1' + \mathbf{R}_2\mathbf{R}_2' + \mathbf{R}_3\mathbf{R}_3'.$$

3. Estimators in the extended growth curve model with a linearly structured covariance matrix

In this section we derive explicit estimators of parameters in the extended growth curve model where the covariance matrix $\mathbf{\Sigma}$ is linearly structured. The linearly structured covariance matrix will be denoted $\mathbf{\Sigma}^{(s)}$. Hence, we consider the extended growth curve model

$$\mathbf{X} = \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2 + \mathbf{E}, \tag{4}$$

deduced from Definition 1.1 for $m = 2$, but with $\mathbf{E} \sim N_{p,n}(\mathbf{0}, \mathbf{\Sigma}^{(s)}, \mathbf{I}_n)$.

The estimation procedure which is proposed in this section will rely on the decompositions of the whole space generated by design matrices as illustrated in Figure 3. The only difference with the unstructured case, Figure 2, being that in Figure 3 we have not replaced $\mathbf{\Sigma}^{(s)}$ with either $\mathbf{S}_1$ nor $\mathbf{S}_2$ because now $\mathbf{\Sigma}^{(s)}$ is structured. If $\mathbf{\Sigma}^{(s)}$ would have been known, we would have a best linear unbiased estimator (BLUE) of the mean

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

where

$\widetilde{\mathbf{M}}_1 = \mathbf{P}_{\mathbf{A}_1,\mathbf{\Sigma}^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_1'}, \widetilde{\mathbf{M}}_2 = \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\mathbf{\Sigma}^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}$ and $\mathbf{T}_1 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\mathbf{\Sigma}^{(s)}}$.

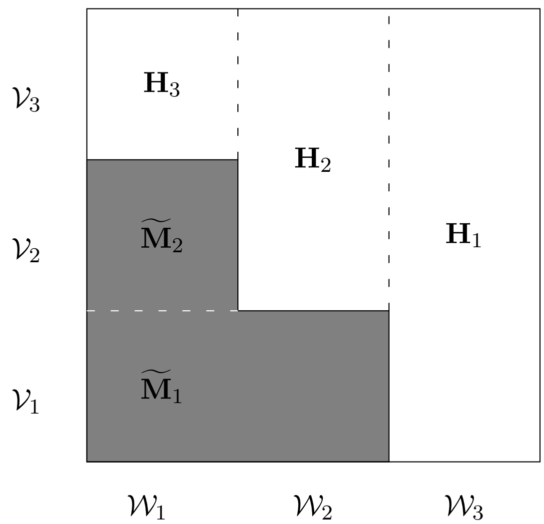

FIGURE 3. Decomposition of the whole space according to the within and between individuals designs illustrating the mean and residual spaces: $\mathcal{V}_1 = \mathcal{C}_{\mathbf{\Sigma}^{(s)}}(\mathbf{A}_1)$, $\mathcal{V}_2 = \mathcal{C}_{\mathbf{\Sigma}^{(s)}}(\mathbf{T}_1\mathbf{A}_2)$, $\mathcal{V}_3 = (\mathcal{C}_{\mathbf{\Sigma}^{(s)}}(\mathbf{A}_1) + \mathcal{C}_{\mathbf{\Sigma}^{(s)}}(\mathbf{T}_1\mathbf{A}_2))^{\perp}$, $\mathcal{W}_1 = (\mathcal{C}(\mathbf{C}_2')$, $\mathcal{W}_2 = (\mathcal{C}(\mathbf{C}_1') \cap \mathcal{C}(\mathbf{C}_2')^{\perp}$, $\mathcal{W}_3 = (\mathcal{C}(\mathbf{C}_1')^{\perp}$.

In Figure 3 we have

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

If $\mathbf{\Sigma}^{(s)}$ is unknown, it should be estimated. The idea is first to estimate $\mathbf{\Sigma}^{(s)}$ when finding the inner product and thereafter reestimate it when it should be interpreted as a dispersion matrix. Now notice that the projection $\mathbf{H}_1$ is independent of $\widetilde{\mathbf{M}}_1$ and that

$$\mathbf{S}_1=\mathbf{H}_1\mathbf{H}_1'=\mathbf{X}(\mathbf{I}-\mathbf{P}_{\mathbf{C}_1'})\mathbf{X}' \sim W_p(\mathbf{\Sigma}^{(s)}, n-r_1) \text{ and } E[\mathbf{S}_1]=(n-r_1)\,\mathbf{\Sigma}^{(s)},$$

where $r_1 = r(\mathbf{C}_1)$. So it is natural to use $\mathbf{S}_1$ when finding inner product estimate. In this paper we apply a least square approach, i.e., we minimize

$$\operatorname{tr}\left\{\left(\mathbf{S}_1-(n-r_1)\mathbf{\Sigma}^{(s)}\right)'\left(\mathbf{S}_1-(n-r_1)\,\mathbf{\Sigma}^{(s)}\right)\right\} \tag{5}$$

with respect to $\mathbf{\Sigma}^{(s)}$.

To find the minimizer of (5), we use techniques based on differentiations. The matrix derivative we use here is defined as follows

Definition 3.1. Let the elements of $\mathbf{Y} \in \mathbb{R}^{r\times s}$ be functions of $\mathbf{X} \in \mathbb{R}^{p\times q}$. The matrix $\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}\mathbf{X}} \in \mathbb{R}^{pq\times rs}$ is called matrix derivative of $\mathbf{Y}$ by $\mathbf{X}$ in a set Ω, if the partial derivatives $\frac{\partial y_{kl}}{\partial x_{ij}}$ exist, are continuous in Ω, and

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}\mathbf{X}}=\frac{\partial}{\partial \mathrm{vec}\mathbf{X}}(\mathrm{vec}'\mathbf{Y}),$$

where

$$\frac{\partial}{\partial \mathrm{vec}\mathbf{X}}=\left(\frac{\partial}{\partial x_{11}},\ldots,\frac{\partial}{\partial x_{p1}},\ldots,\frac{\partial}{\partial x_{1q}},\ldots,\frac{\partial}{\partial x_{pq}}\right)'.$$

It can be shown that the identity

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}\mathbf{X}}=\frac{\mathrm{d}\,\mathrm{vec}'\mathbf{Y}}{\mathrm{d}\,\mathrm{vec}\mathbf{X}}$$

holds. For more details on matrix differentiation one can consult [3].

For convenience we define and denote by $\mathrm{vec}\mathbf{\Sigma}(K)$ the columnwise vectorization of $\mathbf{\Sigma}^{(s)}$ where all 0's and repeated elements (by modulus) have been disregarded. Then there exists, see [3], a transformation matrix $\mathbf{T}$ such that

$$\mathrm{vec}\mathbf{\Sigma}(K)=\mathbf{T}\mathrm{vec}\mathbf{\Sigma}^{(s)} \text{ or } \mathrm{vec}\mathbf{\Sigma}^{(s)}=\mathbf{T}^{+}\mathrm{vec}\mathbf{\Sigma}(K), \tag{6}$$

where $\mathbf{T}^{+}$ denotes the Moore-Penrose generalized inverse of $\mathbf{T}$. Moreover

$$\frac{\mathrm{d}\mathbf{\Sigma}^{(s)}}{\mathrm{d}\mathbf{\Sigma}(K)}=(\mathbf{T}^{+})'. \tag{7}$$

The expression (5) gives

$$\operatorname{tr}\left\{(\mathbf{S}_1-(n-r_1)\mathbf{\Sigma}^{(s)})'(\,)\right\}=(\mathrm{vec}\mathbf{S}_1-(n-r_1)\mathrm{vec}\mathbf{\Sigma}^{(s)})'(\,),$$

where the notation $(\mathbf{Q})'()$ stands for $(\mathbf{Q})'(\mathbf{Q})$.

Differentiating with respect to $\mathrm{vec}\mathbf{\Sigma}(K)$ and equalizing to $\mathbf{0}$, we get

$$-2(n-r_1)\frac{\mathrm{d}\mathbf{\Sigma}^{(s)}}{\mathrm{d}\mathbf{\Sigma}(K)}\mathrm{vec}(\mathbf{S}_1-(n-r_1)\,\mathbf{\Sigma}^{(s)})=\mathbf{0}. \tag{8}$$

From (6), (7) and (8) we obtain the linear equation

$$(\mathbf{T}^+)'\text{vec}\mathbf{S}_1 = (n-r_1)\,(\mathbf{T}^+)'\mathbf{T}^+\text{vec}\boldsymbol{\Sigma}(K),$$

which is consistent and its general solution is given by

$$\text{vec}\boldsymbol{\Sigma}(K) = \frac{1}{n-r_1}\left((\mathbf{T}^+)'\mathbf{T}^+\right)^-(\mathbf{T}^+)'\text{vec}\mathbf{S}_1 + \left((\mathbf{T}^+)'\mathbf{T}^+\right)^o\mathbf{z},$$

where $\mathbf{z}$ is an arbitrary vector. Hence, using (6) we obtain the unique minimizer of (5) given by

$$\text{vec}\boldsymbol{\Sigma}^{(s)} = \mathbf{T}^+\text{vec}\boldsymbol{\Sigma}(K) = \frac{1}{n-r_1}\mathbf{T}^+\left((\mathbf{T}^+)'\mathbf{T}^+\right)^-(\mathbf{T}^+)'\text{vec}\mathbf{S}_1.$$

Thus, a first estimator for $\boldsymbol{\Sigma}^{(s)}$ is given by

$$\text{vec}\widehat{\boldsymbol{\Sigma}}_1^{(s)} = \frac{1}{n-r_1}\mathbf{T}^+\left((\mathbf{T}^+)'\mathbf{T}^+\right)^-(\mathbf{T}^+)'\text{vec}\mathbf{S}_1. \tag{9}$$

Now suppose $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ is positive definite (which always holds for large n). Then using $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ to define the inner product, we may consider $\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_1^{(s)}}(\mathbf{A}_1)$ instead of $\mathcal{C}_{\boldsymbol{\Sigma}^{(s)}}(\mathbf{A}_1)$. Thus an estimator of $\mathbf{M}_1$, and also that of $\mathbf{H}_2$ are found by projecting observations on $\mathcal{C}(\mathbf{C}_1')\otimes\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_1^{(s)}}(\mathbf{A}_1)$ and $\left(\mathcal{C}(\mathbf{C}_1')\cap\mathcal{C}(\mathbf{C}_2')^\perp\right)\otimes\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_1^{(s)}}(\mathbf{A}_1)^\perp$ respectively, i.e.,

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

To derive a second estimator of $\boldsymbol{\Sigma}^{(s)}$ we use a similar idea as above but now we consider the sum of $\mathbf{S}_1$ and $\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'$. Notice that

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

Put $\widehat{\mathbf{T}}_1 = \mathbf{I}-\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}$ and $\mathbf{W}_0 = \mathbf{X}(\mathbf{P}_{\mathbf{C}_1'}-\mathbf{P}_{\mathbf{C}_2'})\mathbf{X}'$. Then $\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' = \widehat{\mathbf{T}}_1\mathbf{W}_0\widehat{\mathbf{T}}_1'$ and $\mathbf{S}_1$ is independent of $\mathbf{W}_0$. Therefore it is natural to condition $\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'$ with respect to $\mathbf{S}_1$ and

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

where $r_2 = r(\mathbf{C}_2')$.

Again we apply a least squares approach and minimize

$$\text{tr}\left\{(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' - [(n-r_1)\,\boldsymbol{\Sigma}^{(s)} + (r_1-r_2)\,\widehat{\mathbf{T}}_1\boldsymbol{\Sigma}^{(s)}\widehat{\mathbf{T}}_1'])'\,(\,)\right\} \tag{11}$$

with respect to $\boldsymbol{\Sigma}^{(s)}$.

The expression (11) gives

$$\operatorname{tr}\left\{(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'-[(n-r_1)\,\boldsymbol{\Sigma}^{(s)}+(r_1-r_2)\,\widehat{\mathbf{T}}_1\boldsymbol{\Sigma}^{(s)}\widehat{\mathbf{T}}_1'])'\,(\,)\right\}$$
$$=(\mathrm{vec}(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2')-\widehat{\boldsymbol{\Upsilon}}\mathrm{vec}\boldsymbol{\Sigma}^{(s)})'(\,) \qquad (12)$$

where

$$\widehat{\boldsymbol{\Upsilon}}=(n-r_1)\mathbf{I}+(r_1-r_2)\widehat{\mathbf{T}}_1\otimes\widehat{\mathbf{T}}_1.$$

Differentiating the expression in righthand side of (12) with respect to $\mathrm{vec}\boldsymbol{\Sigma}(K)$ and equalizing to $\mathbf{0}$, we get

$$-2\frac{\mathrm{d}\boldsymbol{\Sigma}^{(s)}}{\mathrm{d}\boldsymbol{\Sigma}(K)}\widehat{\boldsymbol{\Upsilon}}'(\mathrm{vec}(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2')-\widehat{\boldsymbol{\Upsilon}}\mathrm{vec}\boldsymbol{\Sigma}^{(s)})=\mathbf{0}. \qquad (13)$$

From (6), (7) and (13) we obtain the linear equation

$$(\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\mathbf{T}^+\mathrm{vec}\boldsymbol{\Sigma}(K)=(\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\mathrm{vec}(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'),$$

which is consistent and a general solution is given by

$$\mathrm{vec}\boldsymbol{\Sigma}(K)=\left((\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\mathrm{vec}(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2')+((\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\mathbf{T}^+)^o\mathbf{z},$$

where $\mathbf{z}$ is an arbitrary vector. Hence, using (6) we obtain a second estimator of $\boldsymbol{\Sigma}^{(s)}$ given by

$$\mathrm{vec}\widehat{\boldsymbol{\Sigma}}_2^{(s)}=\mathbf{T}^+\left((\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\mathrm{vec}(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'). \qquad (14)$$

Now using $\widehat{\boldsymbol{\Sigma}}_2^{(s)}$, again assumed to be positive definite, to define the inner product, we may consider $\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_2^{(s)}}(\widehat{\mathbf{T}}_1\mathbf{A}_1)$ instead of $\mathcal{C}_{\boldsymbol{\Sigma}^{(s)}}(\mathbf{T}_1\mathbf{A}_1)$. Thus an estimator of $\mathbf{M}_2$, and also that of $\mathbf{H}_3$ are found by projecting observations on

$$\mathcal{C}(\mathbf{C}_2')\otimes\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_2^{(s)}}(\widehat{\mathbf{T}}_1\mathbf{A}_1)\ \text{ and }\ \mathcal{C}(\mathbf{C}_2')\otimes\left(\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_1^{(s)}}(\mathbf{A}_1)+\mathcal{C}_{\widehat{\boldsymbol{\Sigma}}_2^{(s)}}(\widehat{\mathbf{T}}_1\mathbf{A}_2)\right)^{\perp}$$

respectively, i.e.,

$$\begin{aligned}\widehat{\mathbf{M}}_2&=\mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'},\\ \widehat{\mathbf{H}}_3&=(\mathbf{I}-\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}-\mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}})\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}.\end{aligned} \qquad (15)$$

To derive a final estimator of $\boldsymbol{\Sigma}^{(s)}$, the idea is to use the total sum of squared residuals and proceed similarly as above. So, we now consider the sum of $\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'$ and $\widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3'$. Notice that

$$\widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3'=\widehat{\mathbf{T}}_2\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}\mathbf{X}'\widehat{\mathbf{T}}_2',$$

where $\widehat{\mathbf{T}}_2 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}} - \mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}}$. Then, letting $\mathbf{W}_1 = \mathbf{X}\mathbf{P}_{\mathbf{C}_2'}\mathbf{X}'$, we get $\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' = \widehat{\mathbf{T}}_2\mathbf{W}_1\widehat{\mathbf{T}}_2'$ and $\mathbf{S}_1 + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'$ is independent of $\mathbf{W}_1$. So again it is natural to condition $\widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3'$ with respect to $\mathbf{S}_1 + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'$ and

$$\widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3'|\mathbf{S}_1 + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' \sim W_p\left(\widehat{\mathbf{T}}_2\boldsymbol{\Sigma}^{(s)}\widehat{\mathbf{T}}_2', r_2\right).$$

Again we apply a least squares approach and minimize

$$\operatorname{tr}\left\{(\mathbf{S} - [(n-r_1)\,\boldsymbol{\Sigma}^{(s)} + (r_1 - r_2)\,\widehat{\mathbf{T}}_1\boldsymbol{\Sigma}^{(s)}\widehat{\mathbf{T}}_1' + r_2\widehat{\mathbf{T}}_2\boldsymbol{\Sigma}^{(s)}\widehat{\mathbf{T}}_2'])\,()'\right\}$$
$$= (\text{vec}\mathbf{S} - \widehat{\boldsymbol{\Phi}}\text{vec}\boldsymbol{\Sigma}^{(s)})'\,() \qquad (16)$$

with respect to $\boldsymbol{\Sigma}^{(s)}$, where $\mathbf{S} = \mathbf{S}_1 + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' + \widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3'$ and

$$\widehat{\boldsymbol{\Phi}} = (n - r_1)\,\mathbf{I} + (r_1 - r_2)\,\widehat{\mathbf{T}}_1 \otimes \widehat{\mathbf{T}}_1 + r_2\widehat{\mathbf{T}}_2 \otimes \widehat{\mathbf{T}}_2.$$

We notice that (16) resembles to (12). So using this analogy, we find that the final estimator of $\boldsymbol{\Sigma}^{(s)}$ is given by

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

We suppose $\widehat{\boldsymbol{\Sigma}}^{(s)}$ to be positive definite. The following theorem summarizes what we discussed above and gives the proposed estimators of parameters in the extended growth curve model with a linearly structured covariance matrix.

Theorem 3.2. *Let the extended growth curve model be given by (4). Then*

(1) *The estimator of the structured covariance matrix $\boldsymbol{\Sigma}^{(s)}$ is given by*

$$\text{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} = \mathbf{T}^+\left((\mathbf{T}^+)'\widehat{\boldsymbol{\Phi}}'\widehat{\boldsymbol{\Phi}}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\widehat{\boldsymbol{\Phi}}'\text{vec}\mathbf{S}. \qquad (17)$$

Here, $\widehat{\boldsymbol{\Sigma}}^{(s)}$ is assumed to be positive definite (which always holds for large n).

(2) *The estimated mean is given by*

$$\widehat{E[\mathbf{X}]} = \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_1'} + \mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}. \qquad (18)$$

4. Properties of the proposed estimators

In this section we study some properties, like unbiasedness and consistency, of the estimators we proposed in Section 3. We start with studying properties of estimators for the covariance matrix $\boldsymbol{\Sigma}^{(s)}$ given in (9), (14) and (17).

Lemma 4.1. *The estimator $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ given in (9) is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$, i.e., $\widehat{\boldsymbol{\Sigma}}_1^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$.*

Proof. We know that $\frac{1}{n-r_1}\text{vec}\mathbf{S}_1 \xrightarrow{p} \text{vec}\boldsymbol{\Sigma}^{(s)}$. So, from (6) and (9), we have

$$\begin{aligned}
\text{vec}\widehat{\boldsymbol{\Sigma}}_1^{(s)} &= \frac{1}{n-r_1}\mathbf{T}^+\left((\mathbf{T}^+)'\mathbf{T}^+\right)^-(\mathbf{T}^+)'\text{vec}\mathbf{S}_1\\
&\xrightarrow{p} \mathbf{T}^+\left((\mathbf{T}^+)'\mathbf{T}^+\right)^-(\mathbf{T}^+)'\text{vec}\boldsymbol{\Sigma}^{(s)}\\
&= \mathbf{T}^+\left((\mathbf{T}^+)'\mathbf{T}^+\right)^-(\mathbf{T}^+)'\mathbf{T}^+\text{vec}\boldsymbol{\Sigma}(K)\\
&= \mathbf{T}^+\text{vec}\boldsymbol{\Sigma}(K) = \text{vec}\boldsymbol{\Sigma}^{(s)},
\end{aligned}$$

which concludes the proof. □

Lemma 4.2. *The estimator $\widehat{\boldsymbol{\Sigma}}_2^{(s)}$ given in (14) is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$, i.e., $\widehat{\boldsymbol{\Sigma}}_2^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$.*

Proof. By Lemma 4.1 and Cramer-Slutsky's theorem [1], we have

$$\begin{aligned}
\widehat{\mathbf{T}}_1 &= \mathbf{I}-\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}} \xrightarrow{p} \mathbf{T}_1 = \mathbf{I}-\mathbf{P}_{\mathbf{A}_1,\boldsymbol{\Sigma}^{(s)}},\\
\widehat{\boldsymbol{\Upsilon}} &= (n-r_1)\mathbf{I}+(r_1-r_2)\widehat{\mathbf{T}}_1\otimes\widehat{\mathbf{T}}_1\\
&\xrightarrow{p} \boldsymbol{\Upsilon} = (n-r_1)\mathbf{I}+(r_1-r_2)\mathbf{T}_1\otimes\mathbf{T}_1.
\end{aligned}$$

Hence, from (14), we have

$$\begin{aligned}
\text{vec}\widehat{\boldsymbol{\Sigma}}_2^{(s)} &= \mathbf{T}^+\left((\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\text{vec}(\mathbf{S}_1+\widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2')\\
&\xrightarrow{p} \mathbf{T}^+\left((\mathbf{T}^+)'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\boldsymbol{\Upsilon}'\text{vec}\Big((n-r_1)\boldsymbol{\Sigma}^{(s)}\\
&\quad +(r_1-r_2)\mathbf{T}_1\boldsymbol{\Sigma}^{(s)}\mathbf{T}_1'\Big)\\
&= \mathbf{T}^+\left((\mathbf{T}^+)'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\boldsymbol{\Upsilon}'\left((n-r_1)\mathbf{I}\right.\\
&\quad \left.+(r_1-r_2)\mathbf{T}_1\otimes\mathbf{T}_1\right)\text{vec}\boldsymbol{\Sigma}^{(s)}\\
&= \mathbf{T}^+\left((\mathbf{T}^+)'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\text{vec}\boldsymbol{\Sigma}^{(s)}\\
&= \mathbf{T}^+\left((\mathbf{T}^+)'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\boldsymbol{\Upsilon}'\boldsymbol{\Upsilon}\mathbf{T}^+\text{vec}\boldsymbol{\Sigma}(K)\\
&= \mathbf{T}^+\text{vec}\boldsymbol{\Sigma}(K) = \text{vec}\boldsymbol{\Sigma}^{(s)},
\end{aligned}$$

since $\boldsymbol{\Upsilon}$ has a full rank and thus the proof is complete. □

Theorem 4.3. *The estimator $\widehat{\boldsymbol{\Sigma}}^{(s)}$ given in (17) is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$, i.e., $\widehat{\boldsymbol{\Sigma}}^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$.*

Proof. Using Lemma 4.1, Lemma 4.2, we have

$$
\begin{aligned}
\widehat{\mathbf{T}}_2 &= \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}} - \mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}} \\
&\xrightarrow{p} \mathbf{T}_2 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\boldsymbol{\Sigma}^{(s)}} - \mathbf{P}_{\mathbf{T}_1\mathbf{A}_2,\boldsymbol{\Sigma}^{(s)}}; \\
\widehat{\boldsymbol{\Phi}} &= (n-r_1)\,\mathbf{I} + (r_1-r_2)\,\widehat{\mathbf{T}}_1 \otimes \widehat{\mathbf{T}}_1 + r_2\widehat{\mathbf{T}}_2 \otimes \widehat{\mathbf{T}}_2 \\
&\xrightarrow{p} \boldsymbol{\Phi} = (n-r_1)\,\mathbf{I} + (r_1-r_2)\,\mathbf{T}_1 \otimes \mathbf{T}_1 + r_2\mathbf{T}_2 \otimes \mathbf{T}_2; \\
\mathrm{vec}\mathbf{S} &= \mathrm{vec}(\mathbf{S}_1 + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' + \widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3') \\
&\xrightarrow{p} \mathrm{vec}\left((n-r_1)\boldsymbol{\Sigma}^{(s)} + (r_1-r_2)\,\mathbf{T}_1\boldsymbol{\Sigma}^{(s)}\mathbf{T}_1' + r_2\mathbf{T}_2\boldsymbol{\Sigma}^{(s)}\mathbf{T}_2'\right) \\
&= ((n-r_1)\mathbf{I} + (r_1-r_2)\,\mathbf{T}_1 \otimes \mathbf{T}_1 + r_2\mathbf{T}_2 \otimes \mathbf{T}_2)\,\mathrm{vec}\boldsymbol{\Sigma}^{(s)} \\
&= \boldsymbol{\Phi}\mathrm{vec}\boldsymbol{\Sigma}^{(s)}.
\end{aligned}
$$

Hence, from (17), we have

$$
\begin{aligned}
\mathrm{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} &= \mathbf{T}^+\left((\mathbf{T}^+)'\widehat{\boldsymbol{\Phi}}'\widehat{\boldsymbol{\Phi}}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\widehat{\boldsymbol{\Phi}}'\mathrm{vec}\mathbf{S} \\
&\xrightarrow{p} \mathbf{T}^+\left((\mathbf{T}^+)'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathrm{vec}\boldsymbol{\Sigma}^{(s)} \\
&= \mathbf{T}^+\left((\mathbf{T}^+)'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{T}^+\right)^-(\mathbf{T}^+)'\boldsymbol{\Phi}'\boldsymbol{\Phi}\mathbf{T}^+\mathrm{vec}\boldsymbol{\Sigma}(K) \\
&= \mathbf{T}^+\mathrm{vec}\boldsymbol{\Sigma}(K) = \mathrm{vec}\boldsymbol{\Sigma}^{(s)},
\end{aligned}
$$

since $\boldsymbol{\Phi}$ has a full rank and thus the proof is complete. □

Now we show that the estimator for the mean given in (18) is unbiased.

Theorem 4.4. *Let the estimator* $\widehat{E[\mathbf{X}]}$ *be given in (18), i.e.,*

$$\widehat{E[\mathbf{X}]} = \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_1'} + \mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}.$$

Then $\widehat{E[\mathbf{X}]}$ *is an unbiased estimator of* $E[\mathbf{X}] = \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2$*, i.e.,*

$$E[\widehat{E[\mathbf{X}]}] = \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2.$$

Proof. From (18), we can write

$$\widehat{E[\mathbf{X}]} = \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'} - \mathbf{P}_{\mathbf{C}_2'}) + (\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}} + \mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}})\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}. \tag{19}$$

Let $\mathbf{A}$ be the partitioned matrix $(\mathbf{A}_1 : \mathbf{A}_2)$. Using the fact that

$$\mathcal{C}(\mathbf{A}_1) \boxplus \mathcal{C}(\widehat{\mathbf{T}}_1\mathbf{A}_2) = \mathcal{C}(\mathbf{A}_1) + \mathcal{C}(\mathbf{A}_2),$$

and uniqueness of projectors property, we can write (19) as

$$\widehat{E[\mathbf{X}]} = \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'} - \mathbf{P}_{\mathbf{C}_2'}) + \mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}.$$

Using linearity of expectation, we have

$$E[\widehat{E[\mathbf{X}]}] = E\left[\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'} - \mathbf{P}_{\mathbf{C}_2'})\right] + E\left[\mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}\right].$$

Next,

$$\begin{aligned} E\left[\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'}-\mathbf{P}_{\mathbf{C}_2'})\right] &= E\left[\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\right]E\left[\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'}-\mathbf{P}_{\mathbf{C}_2'})\right] \\ &= E\left[\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\right](\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2)(\mathbf{P}_{\mathbf{C}_1'}-\mathbf{P}_{\mathbf{C}_2'}) \\ &= \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1-\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1\mathbf{P}_{\mathbf{C}_2'}, \end{aligned}$$

where we used the independence of $\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}$ and $\mathbf{X}(\mathbf{P}_{\mathbf{C}_1'}-\mathbf{P}_{\mathbf{C}_2'})$, the fact that $E[\mathbf{X}]=\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2$, $\mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}\mathbf{A}_1=\mathbf{A}_1$, $\mathbf{C}_1\mathbf{P}_{\mathbf{C}_1'}=\mathbf{C}_1$ and $\mathbf{C}_2(\mathbf{P}_{\mathbf{C}_1'}-\mathbf{P}_{\mathbf{C}_2'})=\mathbf{0}$.

Similarly,

$$\begin{aligned} E\left[\mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}\right] &= E\left[\mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\right]E\left[\mathbf{X}\mathbf{P}_{\mathbf{C}_2'}\right] \\ &= E\left[\mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\right](\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2)\mathbf{P}_{\mathbf{C}_2'} \\ &= \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1\mathbf{P}_{\mathbf{C}_2'}+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2, \end{aligned}$$

since $\mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{A}_1=\mathbf{A}_1$, $\mathbf{P}_{\mathbf{A},\widehat{\boldsymbol{\Sigma}}_2^{(s)}}\mathbf{A}_2=\mathbf{A}_2$ and $\mathbf{C}_2\mathbf{P}_{\mathbf{C}_2'}=\mathbf{C}_2$.

Hence,

$$\begin{aligned} E[\widehat{E[\mathbf{X}]}] &= \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1-\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1\mathbf{P}_{\mathbf{C}_2'}+\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1\mathbf{P}_{\mathbf{C}_2'}+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2 \\ &= \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2, \end{aligned}$$

which completes the proof of Theorem 4.4. □

5. Numerical examples

Example 5.1 (Simulation with a banded covariance matrix)**.** Data is generated from $\mathbf{X}\sim N_{p,n}(\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1+\mathbf{A}_2\mathbf{B}_2\mathbf{C}_2,\boldsymbol{\Sigma},\mathbf{I})$, with $p=5$, $n=500$, design matrices are, using the notation $\mathbf{1}_{n/2}$ (respectively $\mathbf{0}_{n/2}$) for a $n/2$-dimension column vector of 1's (respectively of 0's),

$$\mathbf{A}_1'=\begin{pmatrix}1&1&1&1&1\\1&2&3&4&5\end{pmatrix},\quad \mathbf{C}_1=\left(\mathbf{1}_{n/2}'\otimes\begin{pmatrix}1\\0\end{pmatrix}:\mathbf{1}_{n/2}'\otimes\begin{pmatrix}0\\1\end{pmatrix}\right)$$
$$\mathbf{A}_2'=\begin{pmatrix}1^2&2^2&3^2&4^2&5^2\end{pmatrix},\quad \mathbf{C}_2=(\mathbf{0}_{n/2}':\mathbf{1}_{n/2}').$$

The parameter matrices are

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

and the covariance matrix has a banded structure of the form

$$\boldsymbol{\Sigma}^{(s)}=\begin{pmatrix}\alpha&\beta&0&0&0\\\beta&\delta&-\alpha&0&0\\0&-\alpha&\gamma&-\beta&0\\0&0&-\beta&\phi&\alpha\\0&0&0&\alpha&\psi\end{pmatrix},\text{ provided that }\boldsymbol{\Sigma}^{(s)}\text{ is positive definite.}$$

For this structure, the transformation matrix $\mathbf{T}$ in (6) is

$$\mathbf{T}' = \frac{1}{20}\begin{pmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 20 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 20 & 0 & 0 \\ 0 & -5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 20 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 20 \end{pmatrix}$$

The estimates of

$$\mathbf{\Sigma}^{(s)} = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 \\ 1 & 3 & -2 & 0 & 0 \\ 0 & -2 & 4 & -1 & 0 \\ 0 & 0 & -1 & 5 & 2 \\ 0 & 0 & 0 & 2 & 6 \end{pmatrix}$$

based on average estimates from 200 simulations using (9), (14) and (17) are, respectively

$$\widehat{\mathbf{\Sigma}}_1^{(s)} = \begin{pmatrix} 2.0024 & 0.9963 & 0 & 0 & 0 \\ 0.9963 & 2.9974 & -2.0024 & 0 & 0 \\ 0 & -2.0024 & 4.0107 & -0.9963 & 0 \\ 0 & 0 & -0.9963 & 5.0006 & 2.0024 \\ 0 & 0 & 0 & 2.0024 & 5.9980 \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}_2^{(s)} = \begin{pmatrix} 2.0020 & 0.9960 & 0 & 0 & 0 \\ 0.9960 & 2.9973 & -2.0020 & 0 & 0 \\ 0 & -2.0020 & 4.0097 & -0.9960 & 0 \\ 0 & 0 & -0.9960 & 5.0020 & 2.0020 \\ 0 & 0 & 0 & 2.0020 & 5.9964 \end{pmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}^{(s)} = \begin{pmatrix} 2.0059 & 0.9956 & 0 & 0 & 0 \\ 0.9956 & 2.9983 & -2.0059 & 0 & 0 \\ 0 & -2.0059 & 4.0102 & -0.9956 & 0 \\ 0 & 0 & -0.9956 & 5.0041 & 2.0059 \\ 0 & 0 & 0 & 2.0059 & 6.0273 \end{pmatrix}.$$

Example 5.2 (Simulation with a circular Toeplitz covariance matrix)**.** Data is generated from $\mathbf{X} \sim N_{p,n}(\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2, \boldsymbol{\Sigma}, \mathbf{I})$, with $p = 5$, $n \in \{10, 50, 100\}$, design matrices are

$$\mathbf{A}_1' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix}, \quad \mathbf{C}_1 = \left(\mathbf{1}_{n/2}' \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \mathbf{1}_{n/2}' \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$
$$\mathbf{A}_2' = \begin{pmatrix} 1^2 & 2^2 & 3^2 & 4^2 \end{pmatrix}, \quad \mathbf{C}_2 = (\mathbf{0}_{n/2}' : \mathbf{1}_{n/2}').$$

The parameter matrices are

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

and the covariance matrix has a Toeplitz structure

$$\boldsymbol{\Sigma}^{(s)} = \begin{pmatrix} \sigma & \rho_1 & \rho_2 & \rho_1 \\ \rho_1 & \sigma & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & \sigma & \rho_1 \\ \rho_1 & \rho_2 & \rho_1 & \sigma \end{pmatrix}, \text{ provided that } \boldsymbol{\Sigma}^{(s)} \text{ is positive definite.}$$

For this structure, the transformation matrix $\mathbf{T}$ in (6) is

$$\mathbf{T}' = \frac{1}{8}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{pmatrix}.$$

The estimates of

$$\mathbf{\Sigma}^{(s)} = \begin{pmatrix} 1 & 0.5 & 0.25 & 0.5 \\ 0.5 & 1 & 0.5 & 0.25 \\ 0.25 & 0.5 & 1 & 0.5 \\ 0.5 & 0.25 & 0.5 & 1 \end{pmatrix}$$

based on average estimates from 200 simulations using (17) are hereafter. For $n = 10$:

$$\widehat{\mathbf{\Sigma}}^{(s)} = \begin{pmatrix} 0.9929 & 0.4854 & 0.2374 & 0.4854 \\ 0.4854 & 0.9929 & 0.4854 & 0.2374 \\ 0.2374 & 0.4854 & 0.9929 & 0.4854 \\ 0.4854 & 0.2374 & 0.4854 & 0.9929 \end{pmatrix}.$$

For $n = 50$:

$$\widehat{\mathbf{\Sigma}}^{(s)} = \begin{pmatrix} 1.0090 & 0.5095 & 0.2635 & 0.5095 \\ 0.5095 & 1.0090 & 0.5095 & 0.2635 \\ 0.2635 & 0.5095 & 1.0090 & 0.5095 \\ 0.5095 & 0.2635 & 0.5095 & 1.0090 \end{pmatrix}.$$

For $n = 100$:

$$\widehat{\mathbf{\Sigma}}^{(s)} = \begin{pmatrix} 1.0007 & 0.4987 & 0.2502 & 0.4987 \\ 0.4987 & 1.0007 & 0.4987 & 0.2502 \\ 0.2502 & 0.4987 & 1.0007 & 0.4987 \\ 0.4987 & 0.2502 & 0.4987 & 1.0007 \end{pmatrix}.$$

From the above simulated examples, Example 1 and Example 2, we see that the estimates of the linearly structured covariance matrices are in a well agreement with the true covariance matrices. However, through simulations for the structure in Example 1, it was noted that the estimates of the covariance matrix may not be positive definite for small n whereas it is always positive definite for the structure in Example 2. More studies on the positive definiteness of the estimates is of interest.

Example 5.3 (Potthoff and Roy (1964) dental data)**.** The aim of this example is to illustrate the theory developed in this paper with real data set and we do not pretend to carry out data analysis. Dental measurements on eleven girls and sixteen boys at four different ages ($t_1 = 8$, $t_2 = 10$, $t_3 = 12$, and $t_4 = 14$) were taken. Each measurement is the distance, in millimeters, from the center of pituitary to pteryo-maxillary fissure. These data are presented in Table 1.

TABLE 1. Dental data

id	gender	t_1	t_2	t_3	t_4	id	gender	t_1	t_2	t_3	t_4
1	F	21.0	20.0	21.5	23.0	12	M	26.0	25.0	29.0	31.0
2	F	21.0	21.5	24.0	25.5	13	M	21.5	22.5	23.0	26.0
3	F	20.5	24.0	24.5	26.0	14	M	23.0	22.5	24.0	27.0
4	F	23.5	24.5	25.0	26.5	15	M	25.5	27.5	26.5	27.0
5	F	21.5	23.0	22.5	23.5	16	M	20.0	23.5	22.5	26.0
6	F	20.0	21.0	21.0	22.5	17	M	24.5	25.5	27.0	28.5
7	F	21.5	22.5	23.0	25.0	18	M	22.0	22.0	24.5	26.5
8	F	23.0	23.0	23.5	24.0	19	M	24.0	21.5	24.5	25.5
9	F	20.0	21.0	22.0	21.5	20	M	23.0	20.5	31.0	26.0
10	F	16.5	19.0	19.0	19.5	21	M	27.5	28.0	31.0	31.5
11	F	24.5	25.0	28.0	28.0	22	M	23.0	23.0	23.5	25.0
						23	M	21.5	23.5	24.0	28.0
						24	M	17.0	24.5	26.0	29.5
						25	M	22.5	25.5	25.5	26.0
						26	M	23.0	24.5	26.0	30.0
						27	M	22.0	21.5	23.5	25.0

Suppose that for both girls and boys we have a linear growth component but additionally for the boys there also exists a second order polynomial structure. Then, the extended growth curve model with two terms is appropriate to model this data:

$$\mathbf{X} \sim N_{p,n}(\mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2, \mathbf{\Sigma}, \mathbf{I}).$$

In this model, the observation matrix is $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_1, \ldots, \mathbf{x}_{27})$, in which eleven first columns correspond to measurements on girls and sixteen last columns correspond to measurements on boys. The design matrices are

$$\mathbf{A}_1' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 8 & 10 & 12 & 14 \end{pmatrix}, \quad \mathbf{C}_1 = \left(\mathbf{1}_{11}' \otimes \begin{pmatrix}1\\0\end{pmatrix} : \mathbf{1}_{16}' \otimes \begin{pmatrix}0\\1\end{pmatrix}\right)$$
$$\mathbf{A}_2' = \begin{pmatrix} 8^2 & 10^2 & 12^2 & 14^2 \end{pmatrix}, \quad \mathbf{C}_2 = (\mathbf{0}_{11}' : \mathbf{1}_{16}').$$

The matrices $\mathbf{B}_1$ and $\mathbf{B}_2$ are parameter matrices and $\mathbf{\Sigma}$ is the unknown positive definite covariance matrix.

The maximum likelihood estimate for the non-structured covariance matrix computed using (1) is

$$\widehat{\mathbf{\Sigma}}_{ML} = \begin{pmatrix} 5.0272 & 2.5066 & 3.6410 & 2.5099 \\ 2.5066 & 3.8810 & 2.6961 & 3.0712 \\ 3.6410 & 2.6961 & 6.0104 & 3.8253 \\ 2.5099 & 3.0712 & 3.8253 & 4.6164 \end{pmatrix}.$$

Assume that the covariance matrix has a Toeplitz structure, i.e.,

$$\mathbf{\Sigma}^{(s)} = \begin{pmatrix} \sigma & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & \sigma & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & \sigma & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & \sigma \end{pmatrix},$$

the estimate of the structured covariance matrices given by (17) is

$$\widehat{\mathbf{\Sigma}}^{(s)} = \begin{pmatrix} 5.2128 & 3.2953 & 3.6017 & 2.7146 \\ 3.2953 & 5.2128 & 3.2953 & 3.6017 \\ 3.6017 & 3.2953 & 5.2128 & 3.2953 \\ 2.7146 & 3.6017 & 3.2953 & 5.2128 \end{pmatrix},$$

and the MLE computed with Proc Mixed in SAS® [10] is

$$\widehat{\mathbf{\Sigma}}^{(s)}_{\mathrm{ML}} = \begin{pmatrix} 4.9368 & 3.0747 & 3.4559 & 2.2916 \\ 3.0747 & 4.9368 & 3.0747 & 3.4559 \\ 3.4559 & 3.0747 & 4.9368 & 3.0747 \\ 2.2916 & 3.4559 & 3.0747 & 4.9368 \end{pmatrix}.$$

From this example, we see that the proposed estimates are close to the maximum likelihood estimates and the conclusion is that the proposed estimators perform well.

References

[1] Cramer, H. (1946), *Mathematical Methods in Statistics*, Princeton University Press, Princeton.
[2] Khatri, C. G. (1966), *A Note on a MANOVA model applied to problems in growth curve*, Ann. Inst. Statist. Math. **18**, 75–86.
[3] Kollo, T., and von Rosen, D. (2005), *Advanced Multivariate Statistics with Matrices*, Springer, Dordrecht, Netherlands.

[4] Ohlson, M., Andrushchenko, Z., and von Rosen, D. (2011), *Explicit estimators under m-dependence for a multivariate normal distribution*, Ann. Inst. Statist. Math. **63**, 29–42.

[5] Ohlson, M., and von Rosen, D. (2010), *Explicit estimators of parameters in the growth curve model with linearly structured covariance matrices*, J. Multivariate Anal. **101**, 1284–1295.

[6] Olkin, I., and Press, S. (1969), *Testing and estimation for a circular stationary model*, Ann. Math. Statist. **40**(4), 1358–1373.

[7] Pothoff, R. F., and Roy, S. N. (1964), *A generalized multivariate analysis of variance model useful especially for growth curve problems*, Biometrika **51**, 313–326.

[8] Rao, C. R. (1965), *The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves*, Biometrika **52**, 447–458.

[9] von Rosen, D. (1989), *Maximum likelihood estimators in multivariate linear normal models*, J. Multivariate Anal. **31**, 187–200.

[10] SAS Institute Inc. (2008). *SAS/STAT® 9.2 User's Guide: The MIXED Procedure (Book Excerpt)*. Cary, NC: SAS Institute Inc.

[11] Srivastava, M. S., and Khatri, C. G. (1979), *An Introduction to Multivariate Statistics*, North Holland, New York.

[12] Votaw, D. F. (1948), *Testing compound symmetry in a normal multivariate distribution*, Ann. Math. Statist. **19**(4), 447–473.

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

Linköping University, Sweden; National University of Rwanda, Rwanda
E-mail address: joseph.nzabanita@liu.se; jnzabanita@nur.ac.rw

Swedish University of Agricultural Sciences, Sweden
E-mail address: Dietrich.von.Rosen@et.slu.se

Linköping University, Sweden
E-mail address: martin.singull@liu.se