UNIVERSITY OF RWANDA
School of Science and Technology
Department of Applied Mathematics

Cassien HABYARIMANA

ESTIMATION OF PARAMETERS IN THE GROWTH CURVE MODEL WITH A LINEARLY STRUCTURED COVARIANCE MATRIX −A SIMULATION STUDY

A Thesis Submitted in Partial Fulfillment of the Requirements for
the degree of Master of Science in Applied Mathematics
at the University of Rwanda
in Collaboration with Linköping University, Sweden

Supervisory Committee:
Martin Singull, Assistant Professor, Linköping University
Joseph Nzabanita, Lecturer, University of Rwanda

Master Degree Program in Applied Mathematics
Specialization: Statistical Modelling and Actuarial Science

This master degree program thesis is funded by UR/SIDA

December 31, 2013

Declaration

This is to certify that this thesis is my own work, that due reference has been made in the text to all other material used, that it is less than $10,000$ words in length, exclusive of figures, tables and bibliographies, and that it has not been previously submitted for any comparable academic award.

Cassien HABYARIMANA :..
Date :..

Dedication

To My Beloved Wife, My Mother and My Family.

Acknowledgements

I would like to thank my supervisor Dr. Martin Singull and my cosupervisor Mr. Joseph Nzabanita, immensely for helping throughout this theses. I am really grateful to Prof. Verdiana Grace Masanja, Dr. Frodualdd Minani and Dr. Fidele Ndahayo for motivating me throughout my studies and thesis work. My special thanks are due to my colleague, Miss. Byukusenge Beatrice for his help on some MATLAB plots. Most importantly, I want to thank my family and my friends for being with me through thick and thin of my life and being a great motivation.

Abstract

In a study made on estimation of parameters in the extended growth curve model with linearly structured covariance matrix [5] through simulation for some linear structures it was noted that the estimates of the covariance matrix $\mathbf{\Sigma}$ may not be positive definite for small sample sizes whereas it is always positive definite for some other structures. In this master thesis the implementation of the algorithms proposed in [6] and [5] for some known linear structures on the covariance matrix $\mathbf{\Sigma}$ was performed and simulation study for different small sample sizes were repeated many times. For these simulations the percentage of non positive definite estimates were produced and the linear structures that always produce the positive definite estimates were identified and classified from all those linear structures studied.

Key words

Growth curve model, Linearly structured covariance matrix, Explicit estimators, Positive definite matrix.

Nomenclature

UR University of Rwanda
SIDA Swedish International Development Agency
MLE Maximum Likelihood Estimator

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and background

A growth curve is an empirical model of the evolution of a quantity over time. Growth curves are widely used in biology for quantities such as population size, body height or biomass [9]. In mathematical statistics, growth curves are often modeled as being continuous stochastic processes, e.g. as being sample paths that almost surely solve stochastic differential equations [11]. The growth curve model has an important application in many areas such as medicine, pharmacy, natural sciences, social sciences, etc. The simple (classical) growth curve model has been extensively studied over many years and it was introduced in [10], also the extended growth curve model was studied by many authors, for instance [5], [12], [17], [22], etc. The estimation of parameters in the growth curve model (simple and extended), when the covariance matrix has some specific linear structure has been discussed by some authors, for example [5], [6].

In [6], when the simple growth curve model with linearly structured covariance matrix is considered, a suggested estimation procedure gives explicit and consistent estimators of both the mean and the linear structure covariance matrix, and in [5], when the extended growth curve model with two terms and a linearly structured covariance matrix is studied, also a suggested estimation procedure gives explicit and consistent estimators of the linear structure covariance matrix. However, through simulation for some linear structures, it was noted in [5] that the estimates of the covariance matrix may not be positive definite for small sample sizes whereas it is always positive for some other structures.

1.2 Statement of the problem

In this thesis, the problem we deal with, can be formulated in the following way:

Let $\mathbf{X} : p \times n$, $\mathbf{A}_i : p \times q_i$, $\mathbf{B}_i : q_i \times k_i$, $\mathbf{C}_i : k_i \times n$, for $i = 1, 2$, $r(\mathbf{C}_1) + p \leq n$ and $\mathcal{C}(\mathbf{C}_2') \subseteq \mathcal{C}(\mathbf{C}_1')$ where $r(\cdot)$ and $\mathcal{C}(\cdot)$ represent the rank and the column space of a matrix, respectively. The extended growth curve model with two terms and a linearly structured covariance matrix is defined as follows,

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

where the columns of $\mathbf{E}$ are assumed to be independently distributed as a p-variate normal distribution with mean zero and a positive definite covariance matrix $\boldsymbol{\Sigma} = (\sigma_{ij})_{i,j=1}^{p}$ i.e., $\mathbf{E} \sim N_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}, \mathbf{I}_n)$. The covariance matrix sigma has some linear structure, e.g., some elements are equal or similar. The matrices $\mathbf{A}_i$ and $\mathbf{C}_i$ often called design matrices, are known matrices whereas matrices $\mathbf{B}_i$ and $\boldsymbol{\Sigma}$ are unknown parameter matrices.

An estimation procedure that handles linearly structured covariance matrices is proposed by [5]. 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 covariance matrix. This idea was first considered by [6] and is exploited by decomposing the residual space, the orthogonal complement to the design space, into orthogonal subspaces. Studying residuals obtained from projections of observations on these subspaces yields explicit consistent estimator of the covariance matrix. However, through simulation for some linear structures, it was noted in [5] that the estimates of the covariance matrix may not be positive definite for small sample sizes whereas it is always positive definite for some other structures or large sample sizes. Hence, we can ask ourselves, how the problem of non positive definiteness for the estimates of the covariance matrix $\boldsymbol{\Sigma}$ for some linearly structured covariance matrices can be identified?

1.3 Objectives

According to the problem statement, the main objective of this thesis is:

- The implementation of the algorithms proposed by [5] and [6] for some linear structures on the covariance matrix $\boldsymbol{\Sigma}$, and

- from a simulation study try to identify which kind of structures the algorithms will produce positive definite estimates and classify them.

1.4 Outline

This thesis is mainly composed by five chapters: The first chapter comprises the motivation and background of the study, the statement of the problem, the objectives and the outline of the thesis. The second chapter deals with the statistical models that are needed for an easy reading of this thesis. The third chapter deals with different linear structures and the estimators in simple and extended growth curve model with a linearly structured covariance matrix where the algorithms proposed by [5] and [6] are considered. The forth chapter contains our contributions where the numerical simulations and discussions are produced. The last chapter is reserved for the conclusions and suggestions for the further works.

Chapter 2

Statistical models

In this chapter we deal with statistical models that are needed for an easy reading of the next chapters, for instance univariate normal distribution, multivariate normal distribution, matrix normal distribution and growth curve models (simple and extended).

2.1 Univariate normal distribution

Definition 2.1 (Univariate normal distribution) *A random variable x is said to be univariate normally distributed with mean value μ and variance σ^2 if, and only if, its probability density function has the following form:*

$$Prob\,(x \mid \mu, \sigma) = (2\pi\sigma^2)^{-1/2} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}, \qquad -\infty < x < \infty.$$

We will use the notation $x \sim N(\mu, \sigma^2)$.

2.2 Multivariate normal distribution

In probability theory and statistics, the multivariate normal distribution, is a generalization of the one-dimensional (univariate) normal distribution to higher dimensions and it can be defined as follows;

Definition 2.2 (Multivariate normal distribution) *A p-dimensional random vector $\mathbf{x}$ is said to be multivariate normally distributed with a p-dimensional mean vector μ and a $p \times p$ positive definite covariance matrix $\mathbf{\Sigma}$ if, and only if, its joint probability density function has the following form:*

$$Prob\,(\mathbf{x} \mid \mu, \mathbf{\Sigma}) = (2\pi)^{-p/2}\,|\mathbf{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}\,(\mathbf{x}-\mu)'\,\mathbf{\Sigma}^{-1}\,(\mathbf{x}-\mu)\right\}.$$

We will use the notation $\mathbf{x} \sim N_p(\mu, \boldsymbol{\Sigma})$*.*

Maximum likelihood estimators $(MLEs)$

Let $\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_n$ be independent observations from a multivariate normal $N_p(\mu, \boldsymbol{\Sigma})$ and let the observation vectors $\mathbf{x}_i$ for $i = 1, 2, \cdots, n$ be the columns in a matrix $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$. The likelihood function is a function of μ and $\boldsymbol{\Sigma}$ given by

$$L(\mu, \boldsymbol{\Sigma}) = \left((2\pi)^p |\boldsymbol{\Sigma}|\right)^{-n/2} \exp\left\{-\frac{1}{2} tr \left\{\boldsymbol{\Sigma}^{-1} \left(\mathbf{X} - \mu \mathbf{1}_n^{'}\right) \left(\mathbf{X} - \mu \mathbf{1}_n^{'}\right)'\right\}\right\}.$$

The $(MLEs)$ of μ and $\boldsymbol{\Sigma}$ obtained by the maximization of the likelihood function $L(\mu, \boldsymbol{\Sigma})$ with respect to μ and $\boldsymbol{\Sigma}$ are then given by

$$\widehat{\mu}_{MLE} = \frac{1}{n}\mathbf{X}\mathbf{1}_n, \qquad \text{and} \qquad \widehat{\boldsymbol{\Sigma}}_{MLE} = \frac{1}{n}\mathbf{X}\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'\right)\mathbf{X}',$$

where $\mathbf{1}_n$ is the n–dimensional vector of 1s, and $\mathbf{I}_n$ is the $n \times n$ identity matrix (for more details about the MLE see [18]).
Note that, $\mathbf{P}_1 = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n'$ is a projection on the space $\mathcal{C}(\mathbf{1}_n)$ and $(\mathbf{I}_n - \mathbf{P}_1)$ is the projection on the space $\mathcal{C}(\mathbf{1}_n)'$, i.e., $(\mathbf{I}_n - \mathbf{P}_1)\mathbf{1}_n = \mathbf{0}$ which implies that $\widehat{\mu}_{MLE}$ and $\widehat{\boldsymbol{\Sigma}}_{MLE}$ are independent.

2.3 Matrix normal distribution

In statistics, the matrix normal distribution is a probability distribution that is a generalization of the multivariate normal distribution to matrix-valued random variables and is defined as follows;

Definition 2.3 (Matrix normal distribution) *A random* $p \times q$ *matrix* $\mathbf{X}$ *is said to be matrix normally distributed with a mean matrix* $\mathbf{M} : p \times q$ *and two positive definite covariance matrices* $\boldsymbol{\Sigma} : p \times p$ *and* $\boldsymbol{\Psi} : q \times q$ *if, and only if, its joint probability density function has the following form*

$$Prob(\mathbf{X} \mid \mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}) =$$
$$(2\pi)^{-\frac{pq}{2}} |\boldsymbol{\Sigma}|^{-\frac{q}{2}} |\boldsymbol{\Psi}|^{-\frac{p}{2}} \exp\left\{-\frac{1}{2} tr \left\{\boldsymbol{\Sigma}^{-1} (\mathbf{X} - \mathbf{M}) \boldsymbol{\Psi}^{-1} (\mathbf{X} - \mathbf{M})'\right\}\right\}.$$

We will use the notation

$$\mathbf{X} \sim MN_{p,q}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}).$$

The matrix normal is related to the multivariate normal distribution in the following way:

$$\mathbf{X} \sim MN_{p,q}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}) \qquad \text{if and only if} \qquad vec\mathbf{X} \sim N_{pq}(vec\mathbf{M}, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}),$$

where $\otimes$ denotes the Kronecker product and $vec\mathbf{M}$ denotes the vectorization of $\mathbf{M}$.

In this case, the covariance $p \times p$ matrix $\boldsymbol{\Sigma}$ describes the covariances between the rows of $\mathbf{X}$ and the covariance $q \times q$ matrix $\boldsymbol{\Psi}$ describes the covariances between the columns of $\mathbf{X}$, The product $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ takes into account the covariances between columns as well as the covariances between rows of $\mathbf{X}$. Therefore, $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ indicates that the overall covariance consists of the products of the covariances in $\boldsymbol{\Psi}$ and in $\boldsymbol{\Sigma}$, respectively.

Maximum likelihood estimators

Let a random sample of n observation matrices $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ be drawn from the matrix normal distribution $\mathbf{X}$, i.e., $\mathbf{X} \sim MN_{p,q}(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$. The likelihood function is given by the product of the densities evaluated at each observation matrix as it was for the multivariate case. The log-likelihood, ignoring the normalizing factor, is given by

$$\ln L(\mathbf{M}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}) =$$
$$-\frac{nq}{2} \ln |\boldsymbol{\Sigma}| - \frac{np}{2} |\boldsymbol{\Psi}| - \frac{1}{2} \sum_{i=1}^{n} tr \left\{ \boldsymbol{\Sigma}^{-1} (\mathbf{X_i} - \mathbf{M}) \boldsymbol{\Psi}^{-1} (\mathbf{X_i} - \mathbf{M})' \right\}.$$

In estimation of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$ respectively there is a problem that all the parameters are not uniquely defined since for every scalar $k \neq 0$ we have

$$\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma} = k\boldsymbol{\Psi} \otimes \frac{1}{k}\boldsymbol{\Sigma}.$$

The way to obtain the unique estimates has been discussed by [15]. Either set $\Psi_{qq} = 1$ or $\sigma_{pp} = 1$.
The likelihood equations for likelihood estimators under the condition $\Psi_{qq} = 1$, are given by [15] as follows:

$$\begin{aligned}
\widehat{\mathbf{M}} &= \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}_i = \overline{\mathbf{X}}, \\
\widehat{\boldsymbol{\Sigma}} &= \frac{1}{nq} \sum_{i=1}^{n} \left(\mathbf{X}_i - \widehat{\mathbf{M}}\right) \widehat{\boldsymbol{\Psi}}^{-1} \left(\mathbf{X}_i - \widehat{\mathbf{M}}\right)', \\
\widehat{\boldsymbol{\Psi}} &= \frac{1}{np} \sum_{i=1}^{n} \left(\mathbf{X}_i - \widehat{\mathbf{M}}\right)' \widehat{\boldsymbol{\Sigma}}^{-1} \left(\mathbf{X}_i - \widehat{\mathbf{M}}\right).
\end{aligned}$$

There is no explicit solutions to these equations and one must rely on an iterative algorithm like the flip-flop algorithm, for more details read [2] or [15].

2.4 Simple growth curve model

Definition 2.4 *[Growth curve model] Let* $\boldsymbol{X} : p \times n$ *be an observation matrix,* $\boldsymbol{A} : p \times q$ *be a design matrix across individuals,* $\boldsymbol{C} : k \times n$ *be a design matrix within individuals, and* $\boldsymbol{B} : q \times k$ *be an unknown matrix.*
Assume that $q \leq p$, $r + p \leq n$ *and* $n - p - q - 1 > 0$ *where* $r = rank(\boldsymbol{C})$*. The growth curve model is given by*

$$\mathbf{X} = \mathbf{ABC} + \mathbf{E}, \tag{2.1}$$

with $\mathbf{E} \sim \mathbf{MN}_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I_n})$ *where* $\mathbf{\Sigma} = (\sigma_{ij})_{i,j=1}^{p}$ *is an unknown parameter covariance matrix.*

In most applications of the model, p is the number of time points observed on each of the n subjects, $(q-1)$ is the degree of the polynomial, and k is the number of groups.

Maximum likelihood estimators

For the simple growth curve model defined in definition 2.4, when the matrices $\mathbf{A}$ and $\mathbf{C}$ are of full rank, i.e., $rank(\mathbf{A}) = q$ and $rank(\mathbf{C}) = k$, the unique MLE of the mean parameter $\mathbf{B}$, is equal (for more details see for example [3], [4], [13]);

$$\widehat{\mathbf{B}}_{MLE} = \left(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{S}^{-1}\mathbf{X}\mathbf{C}'\left(\mathbf{CC}'\right)^{-1}, \tag{2.2}$$

where $\mathbf{S}$ is the sum of squares matrix given by

$$\mathbf{S} = \mathbf{X}\left(\mathbf{I} - \mathbf{C}'\left(\mathbf{CC}'\right)^{-1}\mathbf{C}\right)\mathbf{X}'. \tag{2.3}$$

Furthermore, the MLE of the covariance matrix $\mathbf{\Sigma}$ is given by

$$n\widehat{\mathbf{\Sigma}}_{MLE} = \left(\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}_{MLE}\mathbf{C}\right)\left(\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}_{MLE}\mathbf{C}\right)' = \widehat{\mathbf{R}}\widehat{\mathbf{R}}' + \widehat{\mathbf{R}}_1\widehat{\mathbf{R}}_1', \tag{2.4}$$

where $\widehat{\mathbf{R}}$ and $\widehat{\mathbf{R}}_1$ are the residuals. By the decomposition of the whole space according to the within and between individual designs illustrating the mean and residual spaces, (see figure 2.1).

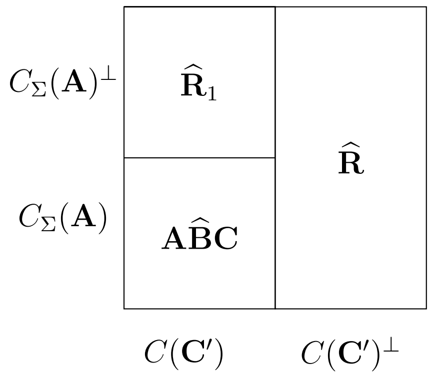

Figure 2.1: Decomposition of the whole space according to the within and between individuals design matrices $\mathbf{A}$ and $\mathbf{C}$.

The residuals $\widehat{\mathbf{R}}$ and $\widehat{\mathbf{R}}_1$ are given by

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

Note that $\mathbf{S} = \widehat{\mathbf{R}}\widehat{\mathbf{R}}'$ does not depend on the mean parameter $\mathbf{B}$ and we know that

$$\frac{1}{n-r}\mathbf{S} \xrightarrow{p} \mathbf{\Sigma}. \tag{2.5}$$

Furthermore, from equation 2.2 it follows that

$$\mathbf{A}\widehat{\mathbf{B}}_{MLE}\mathbf{C} = \mathbf{A}\left(\mathbf{A}'\mathbf{S}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{S}^{-1}\mathbf{X}\mathbf{C}'\left(\mathbf{C}\mathbf{C}'\right)^{-1}\mathbf{C} \tag{2.6}$$

is also unique and therefore $\widehat{\mathbf{\Sigma}}_{MLE}$ is also always uniquely estimated [4].

2.5 Extended growth curve model

In this section we will discuss on the extended growth curve model with only two terms. This model is defined by

Definition 2.5 (Extended growth curve model) *The extended growth curve model with two terms is given by*

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

with $\mathbf{X} : p\times n$ *is an observation matrix,* $\mathbf{A}_i : p\times q_i$ *and* $\mathbf{C}_i : k_i\times n$ *for i=1,2 are known design matrices, and* $\mathbf{B}_i : q_i\times k_i$ *is an unknown parameter matrix where*

$r(\mathbf{C}_1) + p \leq n$ and $\mathcal{C}(\mathbf{C}_2') \subseteq \mathcal{C}(\mathbf{C}_1')$, where $r(\cdot)$ and $\mathcal{C}(\cdot)$ represent the rank and column space of a matrix, respectively. The 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. $\mathbf{E} \sim MN_{p,n}(\mathbf{0}, \mathbf{\Sigma}, \mathbf{I}_n)$. The matrix $\mathbf{\Sigma} = (\sigma_{ij})_{i,j=1}^{p}$ is unknown parameter matrix.

Maximum likelihood estimators

Consider the extended growth curve model defined in definition 2.5. In this case, under general settings, assuming that matrices $\mathbf{A}_i$, $\mathbf{C}_i$ for $i = 1, 2$ are of full rank and that $\mathcal{C}(\mathbf{A}_1) \cap \mathcal{C}(\mathbf{A}_2) = \{0\}$, the unique MLEs $\mathbf{B}_1$ and $\mathbf{B}_2$ are

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

where

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

and the unique estimator of $\mathbf{\Sigma}$ is given by

$$n\widehat{\mathbf{\Sigma}} = \mathbf{R}_1\mathbf{R}_1' + \mathbf{R}_2\mathbf{R}_2' + \mathbf{R}_3\mathbf{R}_3'. \tag{2.7}$$

For any pair of matrices $\mathbf{S}$ and $\mathbf{A}$ where $\mathbf{S}$ is positive definite, we define the following projector (more details see [6]),

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

If $\mathbf{S} = \mathbf{I}$, then $\mathbf{P}_{\mathbf{A},\mathbf{I}}$ is simply noted as $\mathbf{P}_{\mathbf{A}}$.
By decomposition of the whole space according to the within and between individual designs illustrating the mean and residual spaces, (see figure 2.2).

The residuals obtained by projecting data to the subspaces are respectively

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

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

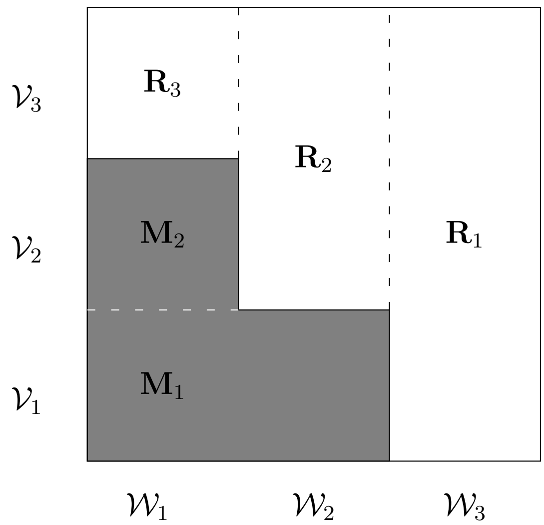

Figure 2.2: Decomposition of the whole space according to the within and between individual designs illustrating the mean and residual spaces.

Chapter 3

Different linear structures for the covariance matrix

In this chapter, we start first by define the linearly structured matrix which is defined as follows;

Definition 3.1 (Linearly structured matrix) *A matrix* $\mathbf{\Sigma} = (\sigma_{ij})$ *is linearly structured if the only linear structure between the elements is given by* $|\sigma_{ij}| = |\sigma_{kl}|$ *and there exists at least one* $(i, j) \neq (k, l)$ *so that* $|\sigma_{ij}| = |\sigma_{kl}|$.

The linear structures for the covariance matrices emerged naturally in statistical applications and they are in the statistical literature for some years ago. These structures are for example, covariance matrix with zeros, banded covariance structure, toeplitz covariance structure, circular toeplitz covariance structure, intraclass covariance structure (also known as uniform covariance structure), compound symmetry structure (type I and II), etc.

The following are some examples of the linearly structured covariance matrices

$$\Sigma_B = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 5 & 2 & 0 \\ 0 & 2 & 6 & 3 \\ 0 & 0 & 3 & 7 \end{pmatrix} \qquad \text{and} \qquad \Sigma_{CT} = \begin{pmatrix} 4 & 1 & 2 & 2 & 1 \\ 1 & 4 & 1 & 2 & 2 \\ 2 & 1 & 4 & 1 & 2 \\ 2 & 2 & 1 & 4 & 1 \\ 1 & 2 & 2 & 1 & 4 \end{pmatrix}.$$

3.1 Covariance matrix with zeros

An example of a covariance matrix with zeros is given by

$$\Sigma_{WZ} = \begin{pmatrix} \sigma_1^2 & 0 & \sigma_{13} & 0 \\ 0 & \sigma_2^2 & 0 & \sigma_{24} \\ \sigma_{13} & 0 & \sigma_3^2 & \sigma_{34} \\ 0 & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{pmatrix}.$$

3.2 Banded covariance structure

The banded covariance structure can for example be

$$\Sigma_B = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & 0 & 0 \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & 0 \\ 0 & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ 0 & 0 & \sigma_{34} & \sigma_4^2 \end{pmatrix}.$$

3.3 Toeplitz and circular toeplitz covariance structure

Toeplitz covariance structure is given by

$$\Sigma_T = \begin{pmatrix} \sigma^2 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & \sigma^2 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & \sigma^2 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & \sigma^2 \end{pmatrix} \quad \text{or} \quad \Sigma_T = \begin{pmatrix} \sigma_1^2 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & \sigma_2^2 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & \sigma_3^2 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & \sigma_4^2 \end{pmatrix}$$

and the circular toeplitz covariance structure is given by

$$\Sigma_{CT} = \begin{pmatrix} \sigma^2 & \rho_1 & \rho_2 & \rho_1 \\ \rho_1 & \sigma^2 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & \sigma^2 & \rho_1 \\ \rho_1 & \rho_2 & \rho_1 & \sigma^2 \end{pmatrix} \quad \text{or} \quad \Sigma_{CT} = \begin{pmatrix} \sigma^2 & \rho_1 & \rho_2 & \rho_2 & \rho_1 \\ \rho_1 & \sigma^2 & \rho_1 & \rho_2 & \rho_2 \\ \rho_2 & \rho_1 & \sigma^2 & \rho_1 & \rho_2 \\ \rho_2 & \rho_2 & \rho_1 & \sigma^2 & \rho_1 \\ \rho_1 & \rho_2 & \rho_2 & \rho_1 & \sigma^2 \end{pmatrix}.$$

3.4 Intraclass (also know as uniform) covariance structure

The intraclass covariance structure looks like

$$\Sigma_{IC} = \begin{pmatrix} \sigma^2 & \rho_0 & \rho_0 & \rho_0 \\ \rho_0 & \sigma^2 & \rho_0 & \rho_0 \\ \rho_0 & \rho_0 & \sigma^2 & \rho_0 \\ \rho_0 & \rho_0 & \rho_0 & \sigma^2 \end{pmatrix}.$$

3.5 Compound symmetry (type I and II)

[20] extended the intraclass model to a model with blocks called compound symmetry, type I and II and are follows

$$\Sigma_{CS-I} = \begin{pmatrix} \alpha & \beta & \beta & \beta \\ \beta & \gamma & \delta & \delta \\ \beta & \delta & \gamma & \delta \\ \beta & \delta & \delta & \gamma \end{pmatrix} \quad \text{and} \quad \Sigma_{CS-II} = \begin{pmatrix} \alpha & \beta & \kappa & \sigma \\ \beta & \alpha & \sigma & \kappa \\ \kappa & \sigma & \gamma & \delta \\ \sigma & \kappa & \delta & \gamma \end{pmatrix}.$$

3.6 Estimators in the growth curve model with a linearly structured covariance matrix

In this section, we tried to develop the estimator of the linearly structured covariance matrix $\mathbf{\Sigma}$ for both classical and extended growth curve models (for more details about estimators in the growth curve model, see [5], [6], [10], [14]).

3.6.1 Estimators in the simple growth curve model with a linearly structured covariance matrix

Consider the following growth curve model defined in Definition 2.4,

$$\mathbf{X} = \mathbf{ABC} + \mathbf{E} \tag{3.1}$$

but with $\mathbf{E} \sim \mathbf{MN}_{p,n}(\mathbf{0}, \mathbf{\Sigma}^{(s)}, \mathbf{I_n})$, where $\mathbf{\Sigma}^{(s)}$ is a linearly structured covariance matrix.

The estimation procedure proposed by [6], when the matrices $\mathbf{A}$ and $\mathbf{C}$ are of full rank, i.e., $rank(\mathbf{A}) = q$ and $rank(\mathbf{C}) = k$, produce the following estimator of the vectorization of the linearly structured covariance matrix $\boldsymbol{\Sigma}^{(s)}$ (for more details read [6])

$$vec\left(\widehat{\boldsymbol{\Sigma}}^{(s)}\right) = \mathbf{T}^{+}\left(\left(\mathbf{T}^{+}\right)'\widehat{\boldsymbol{\Psi}}'\widehat{\boldsymbol{\Psi}}\mathbf{T}^{+}\right)^{-1}\left(\mathbf{T}^{+}\right)'\widehat{\boldsymbol{\Psi}}'vec\left(\mathbf{S}+\widehat{\mathbf{R}}_1\widehat{\mathbf{R}}_1'\right), \quad (3.2)$$

where

$$\widehat{\boldsymbol{\Psi}} = r\left(\mathbf{I}-\mathbf{A}\left(\mathbf{A}'\boldsymbol{\Theta}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\boldsymbol{\Theta}^{-1}\right)\otimes\left(\mathbf{I}-\mathbf{A}\left(\mathbf{A}'\boldsymbol{\Theta}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\boldsymbol{\Theta}^{-1}\right)+(n-r)\mathbf{I}$$

and the residual

$$\widehat{\mathbf{R}}_1 = \left(\mathbf{I}-\mathbf{A}\left(\mathbf{A}'\boldsymbol{\Theta}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\boldsymbol{\Theta}^{-1}\right)\mathbf{X}\mathbf{C}'\left(\mathbf{C}\mathbf{C}'\right)^{-1}\mathbf{C}$$

with $\boldsymbol{\Theta} = \widehat{\boldsymbol{\Sigma}}_1^{(s)}$ is an estimator of $\boldsymbol{\Sigma}^{(s)}$ when the residual $\mathbf{R}_1$ is not considered and is given by

$$vec\widehat{\boldsymbol{\Sigma}}_1^{(s)} = \frac{1}{n-r}\mathbf{T}^{+}\left(\left(\mathbf{T}^{+}\right)'\mathbf{T}^{+}\right)^{-1}\left(\mathbf{T}^{+}\right)'vec\mathbf{S}. \quad (3.3)$$

$\mathbf{T}$ is a matrix such that

$$vec\boldsymbol{\Sigma}^{(s)}(K) = \mathbf{T}vec\boldsymbol{\Sigma}^{(s)}, \quad (3.4)$$

where $vec\boldsymbol{\Sigma}^{(s)}(K)$ is a columnwise vectorized form of $\boldsymbol{\Sigma}^{(s)}$ where all 0 and repeated elements (by absolute value) have been disregarded.

Properties of the proposed estimators

The estimators 3.3, and 3.2 proposed by [6] have some properties (e.g. unbiasedness, and consistency) for more details see [6]:

- The estimator $\widehat{\boldsymbol{\Sigma}}_1^{(p)}$ given in 3.3 is a consistent estimator of $\boldsymbol{\Sigma}^{(p)}$, i.e., $\widehat{\boldsymbol{\Sigma}}_1^{(p)} \xrightarrow{p} \boldsymbol{\Sigma}^{(p)}$, ($\xrightarrow{p}$ denotes convergence in probability).
- The estimator $\widehat{\boldsymbol{\Sigma}}^{(p)}$ given in 3.2 is a consistent estimator of $\boldsymbol{\Sigma}^{(p)}$, i.e., $\widehat{\boldsymbol{\Sigma}}^{(p)} \xrightarrow{p} \boldsymbol{\Sigma}^{(p)}$.

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

Consider again the extended growth curve model given in definition 2.5

$$\mathbf{X} = \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2 + \mathbf{E}, \quad (3.5)$$

but with $\mathbf{E} \sim MN_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}^{(s)}, \mathbf{I}_n)$ where $\boldsymbol{\Sigma}^{(s)}$ is a linearly structured covariance matrix. Assuming that matrices $\mathbf{A}_i$, $\mathbf{C}_i$ for $i = 1, 2$ are of full rank and that $\mathcal{C}(\mathbf{A}_1) \cap \mathcal{C}(\mathbf{A}_2) = \{0\}$.
The main estimator of the structured covariance matrix $\boldsymbol{\Sigma}^{(s)}$ proposed by [5] equals

$$\mathrm{vec}\widehat{\boldsymbol{\Sigma}}^{(s)} = \mathbf{T}^{+}\left((\mathbf{T}^{+})'\widehat{\boldsymbol{\Phi}}'\widehat{\boldsymbol{\Phi}}\mathbf{T}^{+}\right)^{-1}(\mathbf{T}^{+})'\widehat{\boldsymbol{\Phi}}'\mathrm{vec}\mathbf{S}. \tag{3.6}$$

where $\mathbf{S} = \widehat{\mathbf{H}}_1\widehat{\mathbf{H}}_1' + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2' + \widehat{\mathbf{H}}_3\widehat{\mathbf{H}}_3'$ and

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

with $\widehat{\mathbf{T}}_2 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}} - \mathbf{P}_{\widehat{\mathbf{T}}_1\mathbf{A}_2,\widehat{\boldsymbol{\Sigma}}_2^{(s)}}$ and $\widehat{\mathbf{T}}_2 = \mathbf{I} - \mathbf{P}_{\mathbf{A}_1,\widehat{\boldsymbol{\Sigma}}_1^{(s)}}$.
$\widehat{\mathbf{H}}_1$, $\widehat{\mathbf{H}}_2$ and $\widehat{\mathbf{H}}_3$ are projectors given by

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

The above projectors are obtained from from the whole space decomposition according to the within and between individual designs illustrating the mean and residual spaces (see figure 3.1).

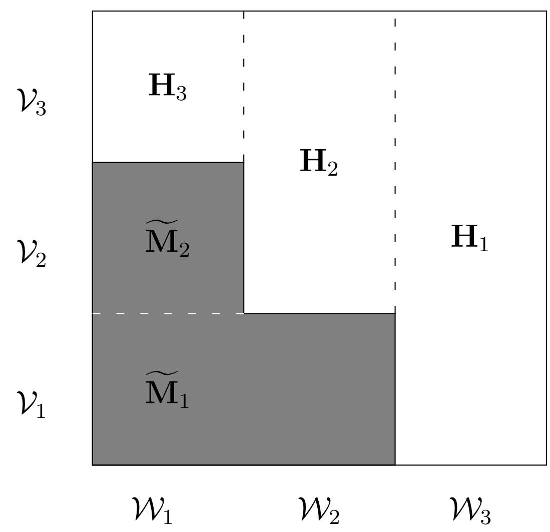

Figure 3.1: Decomposition of the whole space according to the within and between individual designs illustrating the mean and residual spaces.

$\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ and $\widehat{\boldsymbol{\Sigma}}_2^{(s)}$ are the estimators of $\boldsymbol{\Sigma}^{(s)}$ obtained by considering only the residual $\widehat{\mathbf{H}}_1$ and by considering both of $\widehat{\mathbf{H}}_1$ and $\widehat{\mathbf{H}}_2$ respectively (see figure 3.1) and are given by

$$\text{vec}\widehat{\boldsymbol{\Sigma}}_1^{(s)} = \frac{1}{n-r_1}\mathbf{T}^+\left((\mathbf{T}^+)'\mathbf{T}^+\right)^{-1}(\mathbf{T}^+)'\text{vec}\mathbf{S}_1. \tag{3.7}$$

and

$$\text{vec}\widehat{\boldsymbol{\Sigma}}_2^{(s)} = \mathbf{T}^+\left((\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\widehat{\boldsymbol{\Upsilon}}\mathbf{T}^+\right)^{-1}(\mathbf{T}^+)'\widehat{\boldsymbol{\Upsilon}}'\text{vec}(\widehat{\mathbf{H}}_1\widehat{\mathbf{H}}_1' + \widehat{\mathbf{H}}_2\widehat{\mathbf{H}}_2'). \tag{3.8}$$

where

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

Properties of the proposed estimators

The estimators developed in Subsection 3.6.2 have some properties like unbiasedness and consistency, for more details see [5].

- The estimator $\widehat{\boldsymbol{\Sigma}}_1^{(s)}$ given in (3.7) is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$, i.e., $\widehat{\boldsymbol{\Sigma}}_1^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$.
- The estimator $\widehat{\boldsymbol{\Sigma}}_2^{(s)}$ given in (3.8) is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$, i.e., $\widehat{\boldsymbol{\Sigma}}_2^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$.
- The estimator $\widehat{\boldsymbol{\Sigma}}^{(s)}$ given in (3.6) is a consistent estimator of $\boldsymbol{\Sigma}^{(s)}$, i.e., $\widehat{\boldsymbol{\Sigma}}^{(s)} \xrightarrow{p} \boldsymbol{\Sigma}^{(s)}$.

Chapter 4

Simulations and Discussions

In this chapter the simulation study was made for every linearly structured covariance matrix discussed in the previous chapter for both simple and extended growth curve model. In each simulation a sample of observations was randomly generated from a p-variate growth curve model using MATLAB version 7.7.0.471(R2008b).
Samples with the number of observations from $n = 10,20,\cdots,$ (small sample size), $n = 100$ (moderate sample size) to $n = 500$ (large sample size) were considered.
The simulations were repeated $N = 1000$ times and the percentage of non positive definite estimates of the covariance matrices for every sample size is calculated and the graphs are also plotted. For every linearly structured covariance matrix the averaged estimate of the covariance matrix $\mathbf{\Sigma}$ is calculated for $n = 10$ (small sample size) and for $n = 500$ (large sample size) where a comparison is made.

4.1 Simple growth curve model with a linearly structured covariance matrix

Covariance matrix with zeros structure

When the covariance matrix with zeros structure is considered, the averaged estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 16 & 0 & 3 & 0 \\ 0 & 4 & 0 & 5 \\ 3 & 0 & 9 & 2 \\ 0 & 5 & 2 & 10 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 16.0813 & 0 & 2.8160 & 0 \\ 0 & 4.0353 & 0 & 4.7343 \\ 2.8160 & 0 & 9.0895 & 2.0610 \\ 0 & 4.7343 & 2.0610 & 10.1970 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 15.9450 & 0 & 2.9820 & 0 \\ 0 & 3.9964 & 0 & 4.9843 \\ 2.9820 & 0 & 8.9819 & 1.9700 \\ 0 & 4.9843 & 1.9700 & 9.9580 \end{pmatrix}$$

The above estimates are closely to the covariance matrix $\boldsymbol{\Sigma}$
The covariance matrix with zeros structure shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for small sample size n and this percentage decreases as well as n increases and we observe that for n greater than 80 this percentage is zero (see figure 4.1 and table 4.1).

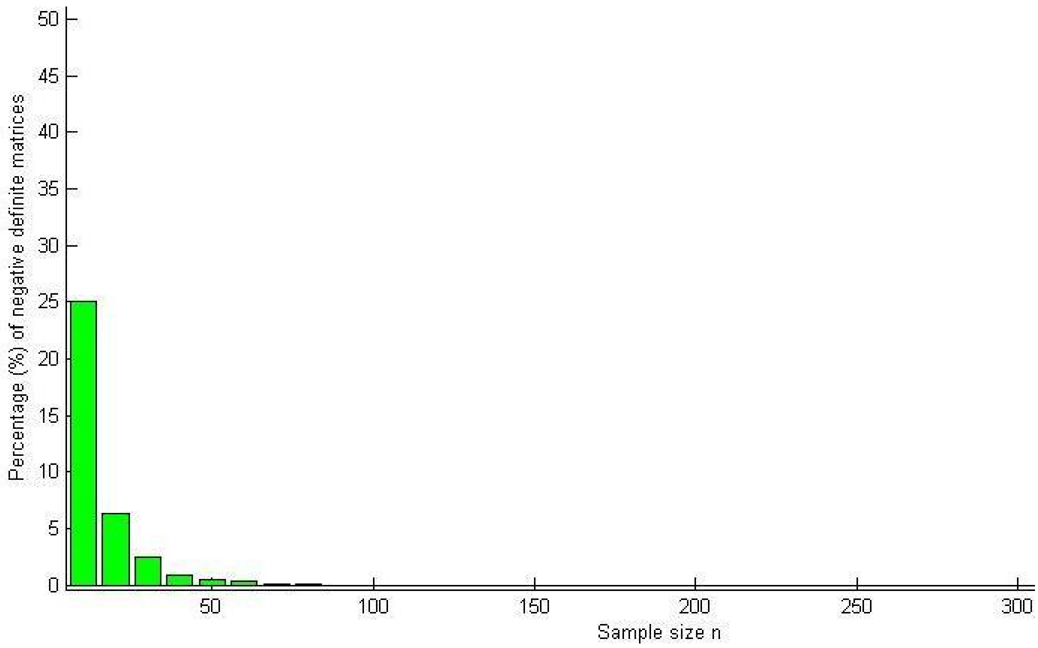

Figure 4.1: Covariance matrix with zeros

Banded covariance structure

When the banded covariance structure (with $p = 4$) is considered, the averaged estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 7 & 1 & 0 & 0 \\ 1 & 4 & 2 & 0 \\ 0 & 2 & 5 & 3 \\ 0 & 0 & 3 & 6 \end{pmatrix}$$

is given ($n = 10$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 7.1165 & 0.9180 & 0 & 0 \\ 0.9180 & 3.9835 & 1.7814 & 0 \\ 0 & 1.7814 & 4.6668 & 2.6926 \\ 0 & 0 & 2.6926 & 5.8630 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 7.0191 & 1.0076 & 0 & 0 \\ 1.0076 & 3.9994 & 1.9921 & 0 \\ 0 & 1.9921 & 4.9887 & 2.9927 \\ 0 & 0 & 2.9927 & 6.0086 \end{pmatrix}$$

In this case, the estimates of $\mathbf{\Sigma}$ are closely to the true value.
The banded covariance structure with $p = 4$ also shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for small sample size n and this percentage decreases as well as n increases and we observe that for n greater than 60 this percentage is zero (see figure 4.2and table 4.1).

Toeplitz covariance structure with the same variances

When the toeplitz covariance structure (with the same variances) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 3 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 4 & 1 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

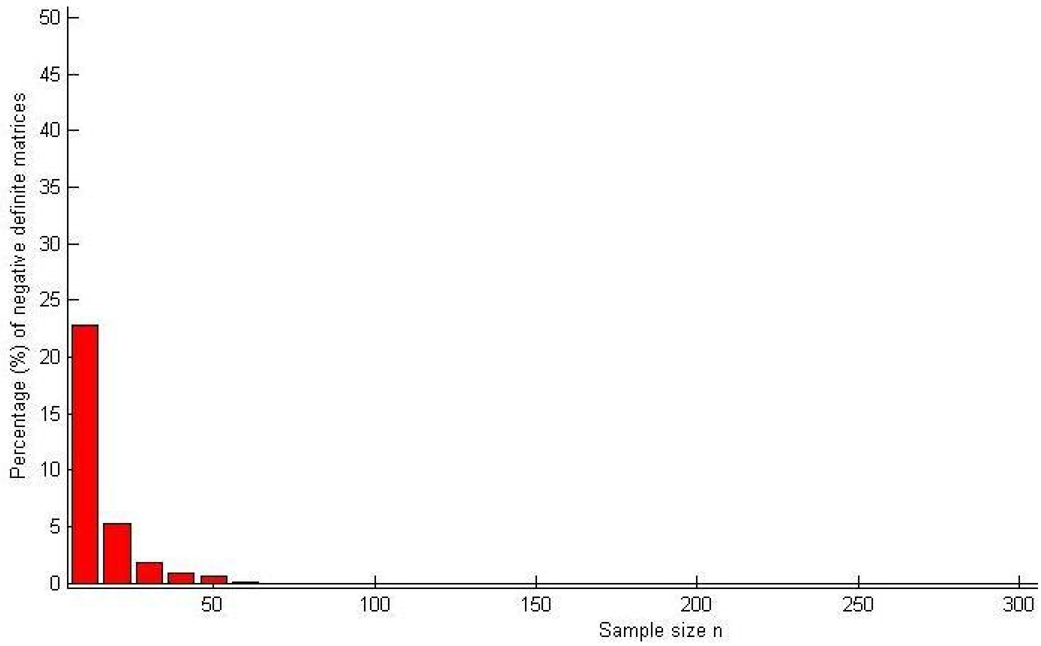

Figure 4.2: Banded covariance structure with $p = 4$

is given by $(n = 10)$

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 3.9638 & 1.0070 & 1.9702 & 2.9270 \\ 1.0070 & 3.9638 & 1.0070 & 1.9702 \\ 1.9702 & 1.0070 & 3.9638 & 1.0070 \\ 2.9270 & 1.9702 & 1.0070 & 3.9638 \end{pmatrix}$$

and $(n = 500)$

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 3.9985 & 1.0050 & 2.0065 & 2.9969 \\ 1.0050 & 3.9985 & 1.0050 & 2.0065 \\ 2.0065 & 1.0050 & 3.9985 & 1.0050 \\ 2.9969 & 2.0065 & 1.0050 & 3.9985 \end{pmatrix}$$

The above estimates of $\boldsymbol{\Sigma}$ are accurate.
The toeplitz covariance structure with the same variances shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for small and moderate sample sizes n but this percentage decreases as well as n increases and also we observe that until $n = 260$ this percentage is greater than zero (see figure 4.3 and table 4.1).

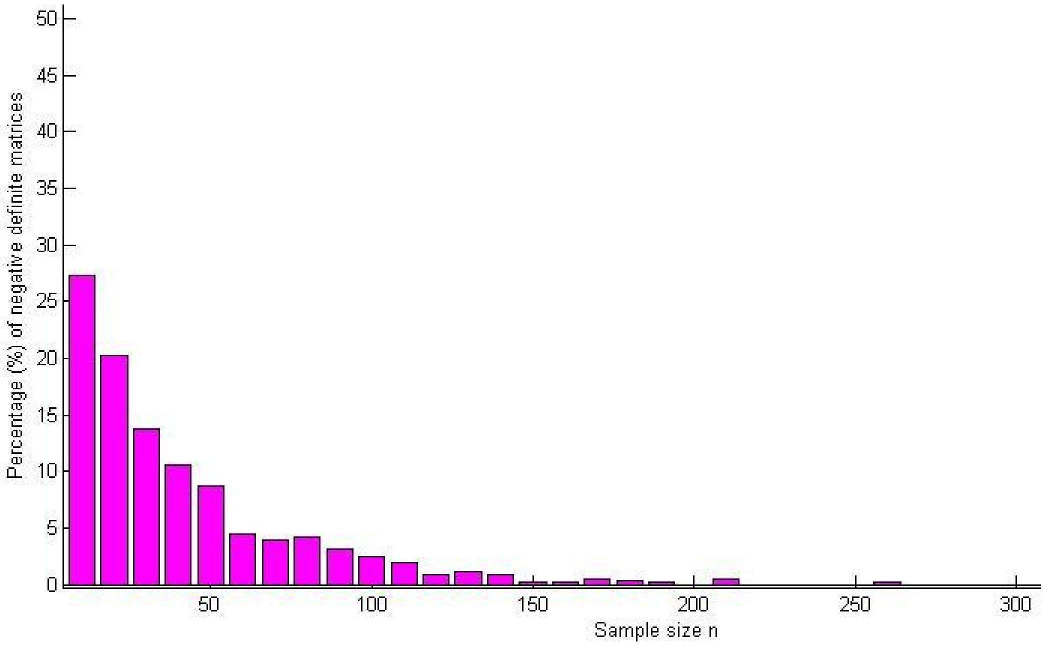

Figure 4.3: Toeplitz covariance structure with same variances

Toeplitz covariance structure with different variances

When the toeplitz covariance structure (with different variances) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 3 \\ 1 & 5 & 1 & 2 \\ 2 & 1 & 6 & 1 \\ 3 & 2 & 1 & 7 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 4.0685 & 0.9443 & 2.0036 & 3.0107 \\ 0.9443 & 5.0873 & 0.9443 & 2.0036 \\ 2.0036 & 0.9443 & 6.0001 & 0.9443 \\ 3.0107 & 2.0036 & 0.9443 & 7.1258 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 4.0047 & 1.0006 & 1.9971 & 3.0057 \\ 1.0006 & 4.9980 & 1.0006 & 1.9971 \\ 1.9971 & 1.0006 & 6.0180 & 1.0006 \\ 3.0057 & 1.9971 & 1.0006 & 7.0138 \end{pmatrix}$$

The above estimates of $\mathbf{\Sigma}$ are accurate.
The toeplitz covariance structure with different variances also shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for small sample size n less than 50 and this percentage decreases as well as n increases and we observe that for n greater than 50 this percentage is zero (see figure 4.4 and table 4.1).

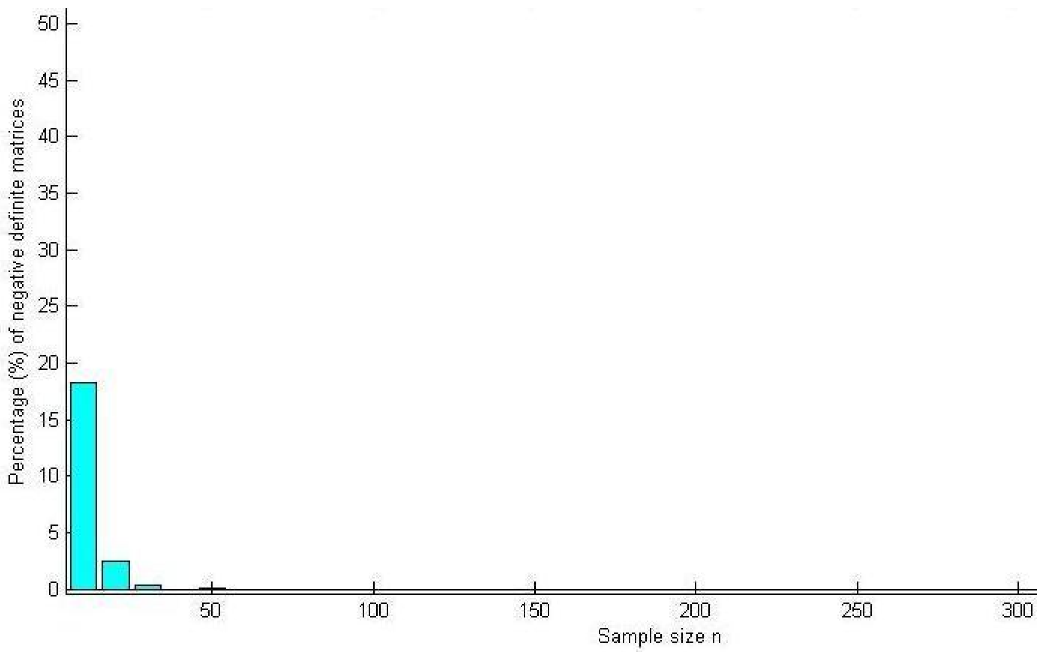

Figure 4.4: Toeplitz covariance structure with different variances

Circular toeplitz covariance structure with $p = 4$

When the circular toeplitz covariance structure (with $p = 4$) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 1 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 4 & 1 \\ 1 & 2 & 1 & 4 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 3.9833 & 1.0108 & 1.9840 & 1.0108 \\ 1.0108 & 3.9833 & 1.0108 & 1.9840 \\ 1.9840 & 1.0108 & 3.9833 & 1.0108 \\ 1.0108 & 1.9840 & 1.0108 & 3.9833 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 4.0026 & 1.0013 & 1.9998 & 1.0013 \\ 1.0013 & 4.0026 & 1.0013 & 1.9998 \\ 1.9998 & 1.0013 & 4.0026 & 1.0013 \\ 1.0013 & 1.9998 & 1.0013 & 4.0026 \end{pmatrix}$$

The above estimates of $\boldsymbol{\Sigma}$ are accurate.
The circular toeplitz covariance structure with $p = 4$ shows a zero percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for all sample sizes n considered (see table 4.1).

Circular toeplitz covariance structure with $p = 5$

When the circular toeplitz covariance structure (with $p = 5$) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 2 & 1 \\ 1 & 4 & 1 & 2 & 2 \\ 2 & 1 & 4 & 1 & 2 \\ 2 & 2 & 1 & 4 & 1 \\ 1 & 2 & 2 & 1 & 4 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 4.0227 & 1.0705 & 2.0549 & 2.0549 & 1.0705 \\ 1.0705 & 4.0227 & 1.0705 & 2.0549 & 2.0549 \\ 2.0549 & 1.0705 & 4.0227 & 1.0705 & 2.0549 \\ 2.0549 & 2.0549 & 1.0705 & 4.0227 & 1.0705 \\ 1.0705 & 2.0549 & 2.0549 & 1.0705 & 4.0227 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 4.0044 & 1.0055 & 2.0031 & 2.0031 & 1.0055 \\ 1.0055 & 4.0044 & 1.0055 & 2.0031 & 2.0031 \\ 2.0031 & 1.0055 & 4.0044 & 1.0055 & 2.0031 \\ 2.0031 & 2.0031 & 1.0055 & 4.0044 & 1.0055 \\ 1.0055 & 2.0031 & 2.0031 & 1.0055 & 4.0044 \end{pmatrix}$$

The above estimates of $\mathbf{\Sigma}$ are accurate.
The circular toeplitz covariance structure also shows a zero percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for all sample size n (see table 4.1).

Intraclass covariance structure

When the intraclass covariance structure is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 1.9665 & 0.9618 & 0.9618 & 0.9618 \\ 0.9618 & 1.9665 & 0.9618 & 0.9618 \\ 0.9618 & 0.9618 & 1.9665 & 0.9618 \\ 0.9618 & 0.9618 & 0.9618 & 1.9665 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 1.9987 & 0.9988 & 0.9988 & 0.9988 \\ 0.9988 & 1.9987 & 0.9988 & 0.9988 \\ 0.9988 & 0.9988 & 1.9987 & 0.9988 \\ 0.9988 & 0.9988 & 0.9988 & 1.9987 \end{pmatrix}$$

The above estimates of $\mathbf{\Sigma}$ are accurate.
The intraclass covariance structure also shows zero percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for all sample size n considered (see table 4.1).

Compound symmetric type I structure

When the compound symmetric type I structure is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 10 & 9 & 9 \\ 1 & 9 & 10 & 9 \\ 1 & 9 & 9 & 10 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 1.9902 & 1.0051 & 1.0051 & 1.0051 \\ 1.0051 & 9.9616 & 8.9580 & 8.9580 \\ 1.0051 & 8.9580 & 9.9616 & 8.9580 \\ 1.0051 & 8.9580 & 8.9580 & 9.9616 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 2.0010 & 0.9967 & 0.9967 & 0.9967 \\ 0.9967 & 10.0392 & 9.0369 & 9.0369 \\ 0.9967 & 9.0369 & 10.0392 & 9.0369 \\ 0.9967 & 9.0369 & 9.0369 & 10.0392 \end{pmatrix}$$

The above estimates of $\boldsymbol{\Sigma}$ are accurate.
The compound symmetric type I structure shows a very small percentage closed to zero of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for small sample size n around 10 and for sample sizes greater than 10, the percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ is zero (see table 4.1).

Compound symmetric type II structure

When the compound symmetric type II structure is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 5 & 2 & 3 & 1 \\ 2 & 5 & 1 & 3 \\ 3 & 1 & 6 & 4 \\ 1 & 3 & 4 & 6 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 4.9712 & 2.0527 & 3.0154 & 1.0704 \\ 2.0527 & 4.9712 & 1.0704 & 3.0154 \\ 3.0154 & 1.0704 & 6.2224 & 4.2368 \\ 1.0704 & 3.0154 & 4.2368 & 6.2224 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 5.0068 & 2.0086 & 3.0044 & 1.0042 \\ 2.0086 & 5.0068 & 1.0042 & 3.0044 \\ 3.0044 & 1.0042 & 6.0120 & 4.0104 \\ 1.0042 & 3.0044 & 4.0104 & 6.0120 \end{pmatrix}$$

The above estimates of $\mathbf{\Sigma}$ are accurate.
The compound symmetric type II structure also shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for small sample size n less than 30 but this percentage decreases as well as n increases and we observe that for n greater than 30 this percentage is zero(see figure 4.5 and table 4.1).

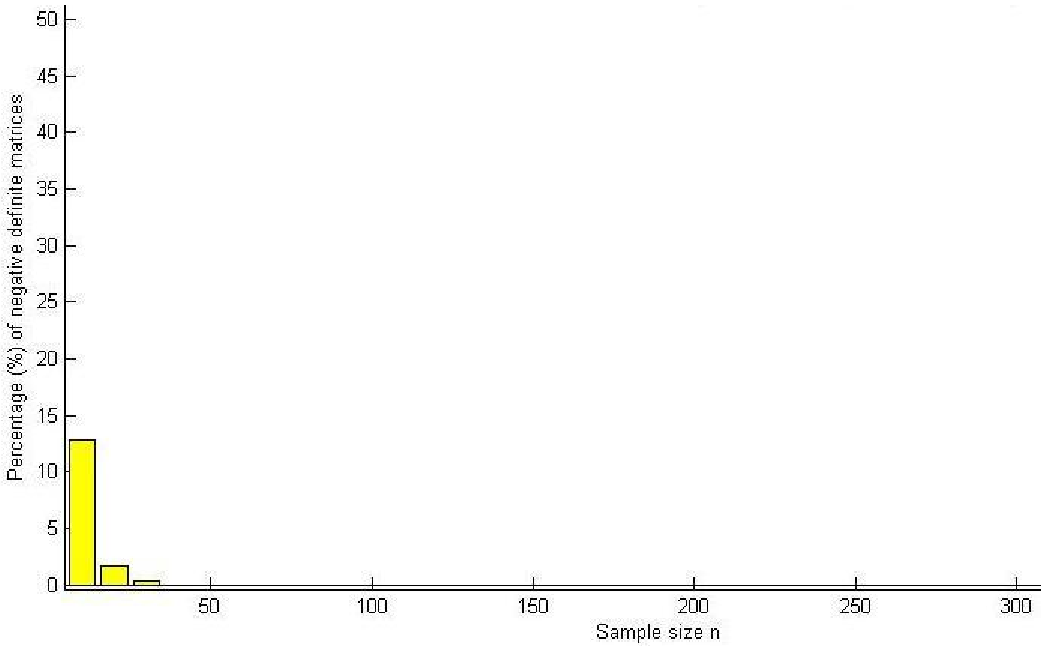

Figure 4.5: Compound symmetric type II

The table (4.1) shows the percentage of non positive definite estimates of the linearly structured covariance matrix $\mathbf{\Sigma}$ for the different sample sizes and linear structures for the simple growth curve model (SGCM) where **S1** stands for the covariance matrix with zeros, **S2** for the banded covariance structure, **S3** for the toeplitz covariance structure with the same variances, **S4** for the toeplitz covariance structure with the different variances, **S5** for the circular toeplitz covariance structure with $p = 4$, **S6** for the circular toeplitz covariance structure with $p = 5$, **S7** for intraclass covariance structure (or uniform covariance structure), and **S8** and **S9** respectively for the compound symmetry structure type I and II.

Table 4.1: Percentage of non positive definite estimates of $\mathbf{\Sigma}$ for the different sample sizes and linear structures for SGCM

Sample size n	Percentage of non positive definite estimates of $\mathbf{\Sigma}$								
	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**
10	25.1	22.8	27.3	18.2	0	0	0	0.1	12.8
20	6.3	5.3	20.3	2.5	0	0	0	0	1.7
30	2.4	1.8	13.7	0.3	0	0	0	0	0.3
40	0.8	0.8	10.5	0	0	0	0	0	0
50	0.4	0.6	8.7	0.1	0	0	0	0	0
60	0.3	0.1	4.4	0	0	0	0	0	0
70	0.1	0	3.9	0	0	0	0	0	0
80	0.1	0	4.2	0	0	0	0	0	0
90	0	0	3.1	0	0	0	0	0	0
100	0	0	2.4	0	0	0	0	0	0
110	0	0	1.9	0	0	0	0	0	0
120	0	0	0.8	0	0	0	0	0	0
130	0	0	1.1	0	0	0	0	0	0
140	0	0	0.8	0	0	0	0	0	0
150	0	0	0.2	0	0	0	0	0	0
160	0	0	0.2	0	0	0	0	0	0
170	0	0	0.4	0	0	0	0	0	0
180	0	0	0.3	0	0	0	0	0	0
190	0	0	0.2	0	0	0	0	0	0
200	0	0	0	0	0	0	0	0	0
210	0	0	0.4	0	0	0	0	0	0
220	0	0	0	0	0	0	0	0	0
230	0	0	0	0	0	0	0	0	0
240	0	0	0	0	0	0	0	0	0
250	0	0	0	0	0	0	0	0	0
260	0	0	0.2	0	0	0	0	0	0
270	0	0	0	0	0	0	0	0	0
280	0	0	0	0	0	0	0	0	0
290	0	0	0	0	0	0	0	0	0
300	0	0	0	0	0	0	0	0	0

4.2 Extended growth curve model with a linearly structured covariance matrix

The same discussions will be made in this section as it was done in the previous one.

Covariance matrix with zeros structure

When the covariance matrix with zeros structure is considered, the averaged estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 16 & 0 & 3 & 0 \\ 0 & 4 & 0 & 5 \\ 3 & 0 & 9 & 2 \\ 0 & 5 & 2 & 10 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 257.0860 & 0 & -88.0042 & 0 \\ 0 & 44.6585 & 0 & -33.2569 \\ -88.0042 & 0 & 64.6777 & 2.7536 \\ 0 & -33.2569 & 2.7536 & 64.7439 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 15.9465 & 0 & 2.9813 & 0 \\ 0 & 3.9964 & 0 & 4.9844 \\ 2.9813 & 0 & 8.9822 & 1.9698 \\ 0 & 4.9844 & 1.9698 & 9.9582 \end{pmatrix}$$

For the above two estimates of $\boldsymbol{\Sigma}$, for small sample size $n = 10$ the estimate is not closed to the proposed value of $\boldsymbol{\Sigma}$.
The covariance matrix with zeros structure shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for small sample size n and this percentage decreases as well as n increases and we observe that for n greater than 60 this percentage is zero (see figure 4.6 and table 4.2).

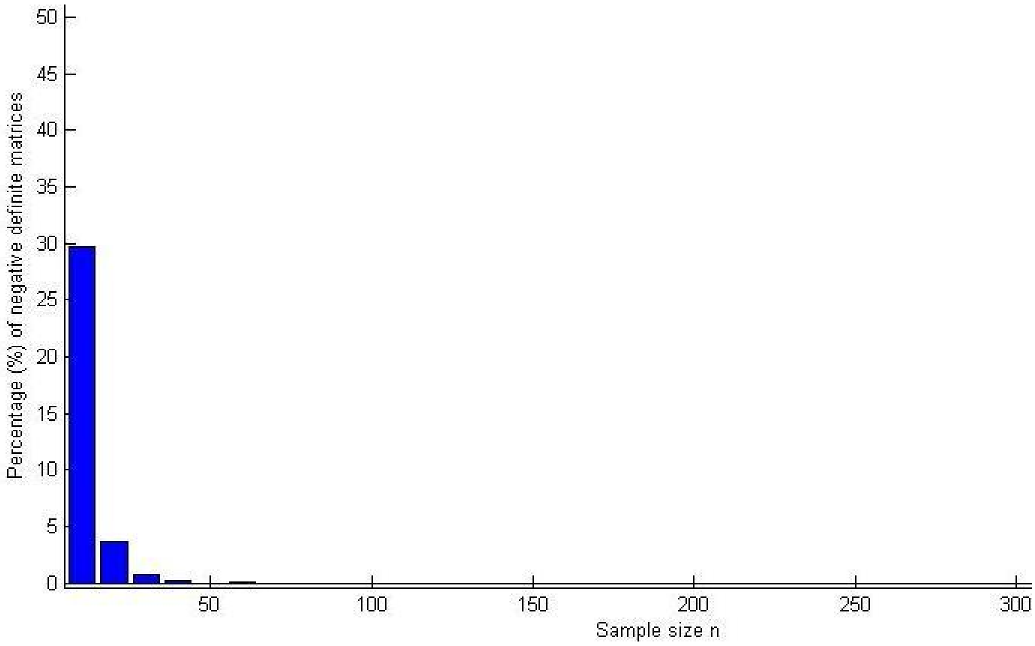

Figure 4.6: Covariance matrix with zeros

Banded covariance structure (with $p = 4$)

When the banded covariance structure (with $p = 4$) is considered, the averaged estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 7 & 1 & 0 & 0 \\ 1 & 4 & 2 & 0 \\ 0 & 2 & 5 & 3 \\ 0 & 0 & 3 & 6 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 77.8090 & -15.6525 & 0 & 0 \\ -15.6526 & 54.3274 & 86.7390 & 0 \\ 0 & 86.7390 & 105.2802 & -16.9658 \\ 0 & 0 & -16.9658 & 114.3014 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 7.0020 & 1.0001 & 0 & 0 \\ 1.0001 & 3.9993 & 2.0070 & 0 \\ 0 & 2.0070 & 5.0035 & 2.9933 \\ 0 & 0 & 2.9933 & 6.0088 \end{pmatrix}$$

For the above two estimates of $\mathbf{\Sigma}$, for small sample size $n = 10$ the estimate is not closed to the proposed value of $\mathbf{\Sigma}$.
The banded covariance structure with $p = 4$ also shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for small sample size n and this percentage decreases as well as n increases and we observe that for n greater than 60 this percentage is zero (see figure 4.7 and table 4.2).

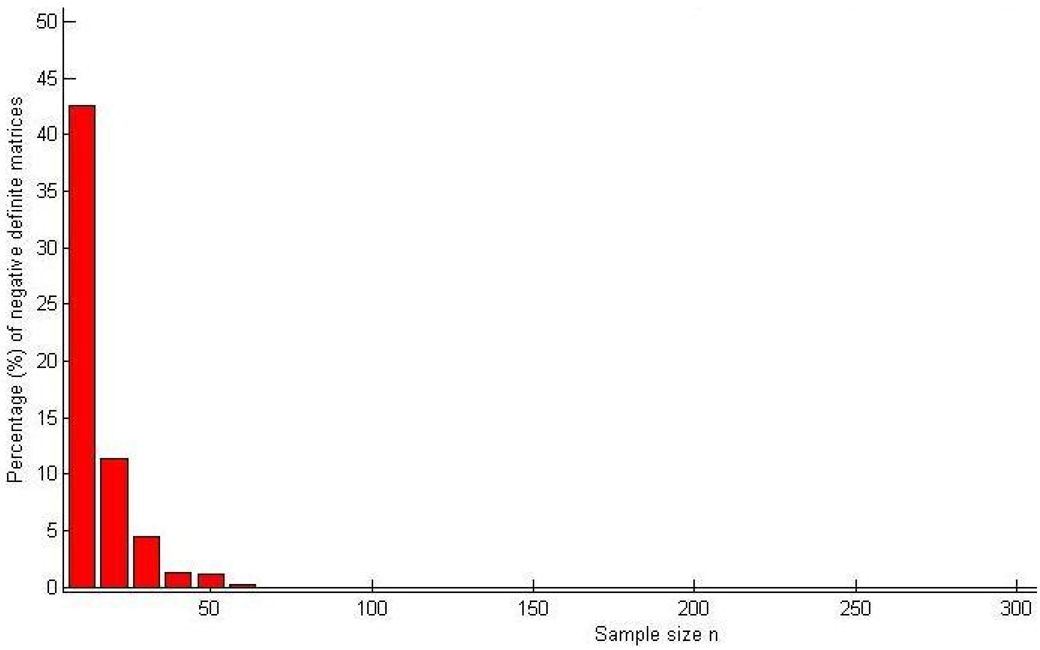

Figure 4.7: Banded covariance structure with p=4

Toeplitz covariance structure (with the same variances)

When the toeplitz covariance structure (with the same variances) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 3 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 4 & 1 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 4.4566 & 1.3752 & 2.1237 & 3.3081 \\ 1.3752 & 4.4566 & 1.3752 & 2.1237 \\ 2.1237 & 1.3752 & 4.4566 & 1.3752 \\ 3.3081 & 2.1237 & 1.3752 & 4.4566 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 3.9884 & 0.9920 & 1.9929 & 2.9854 \\ 0.9920 & 3.9884 & 0.9920 & 1.9929 \\ 1.9929 & 0.9920 & 3.9884 & 0.9920 \\ 2.9854 & 1.9929 & 0.9920 & 3.9884 \end{pmatrix}$$

In this case, the estimates of $\mathbf{\Sigma}$ are closely to the true value.
The toeplitz covariance structure with the same variances shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for small and moderate sample sizes n but this percentage decreases as well as n increases and we observe that until $n = 300$ this percentage is greater than zero (see figure 4.8 and table 4.2).

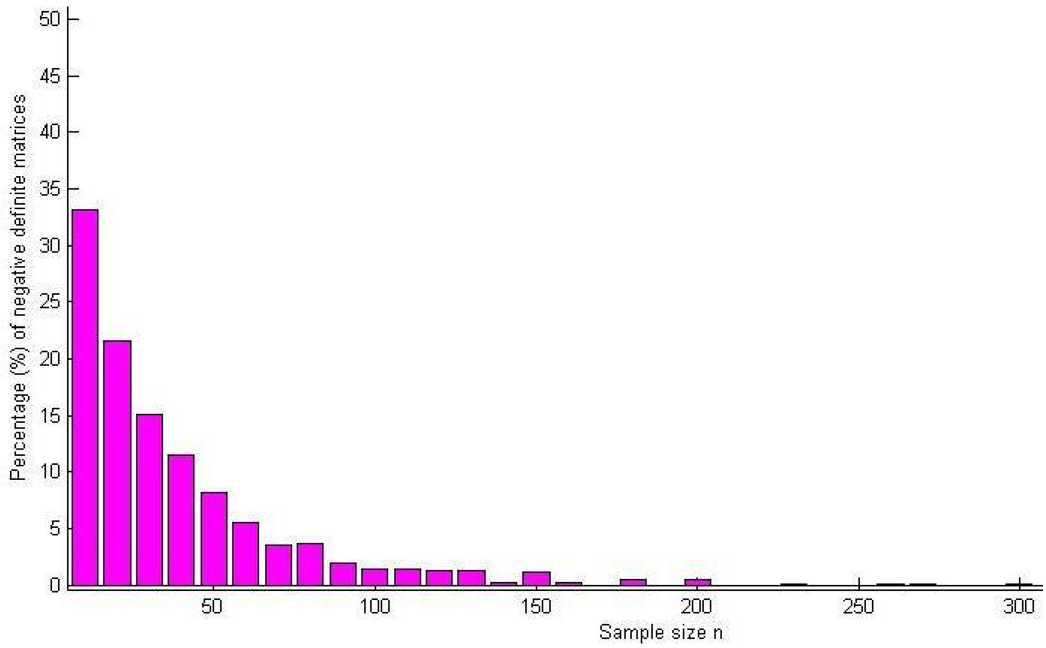

Figure 4.8: Toeplitz covariance structure with same variances

Toeplitz covariance structure (with different variances)

When the toeplitz covariance structure (with different variances) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 3 \\ 1 & 5 & 1 & 2 \\ 2 & 1 & 6 & 1 \\ 3 & 2 & 1 & 7 \end{pmatrix}$$

is given by $(n = 10)$

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 123.9509 & 0.1276 & -28.6401 & 83.8243 \\ 0.1276 & 30.0239 & 0.1276 & -28.6401 \\ -28.6401 & 0.1276 & 22.8943 & 0.1276 \\ 83.8243 & -28.6401 & 0.1276 & 149.9859 \end{pmatrix}$$

and $(n = 500)$

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 3.9864 & 0.9883 & 1.9890 & 2.9801 \\ 0.9883 & 4.9965 & 0.9883 & 1.9890 \\ 1.9890 & 0.9883 & 5.9832 & 0.9883 \\ 2.9801 & 1.9890 & 0.9883 & 6.9681 \end{pmatrix}$$

In this case, the estimates of $\boldsymbol{\Sigma}$ are closely to the true value.
For the above two estimates of $\boldsymbol{\Sigma}$, for small sample size $n = 10$ the estimate is not closed to the proposed value of $\boldsymbol{\Sigma}$.
The toeplitz covariance structure with different variances also shows a big percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for small sample size n less than 40 and this percentage decreases as well as n increases and we observe that for n greater than 40 this percentage is zero(see figure 4.9 and table 4.2).

Circular toeplitz covariance structure (with $p = 4$)

When the circular toeplitz covariance structure (with $p = 4$) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 1 \\ 1 & 4 & 1 & 2 \\ 2 & 1 & 4 & 1 \\ 1 & 2 & 1 & 4 \end{pmatrix}$$

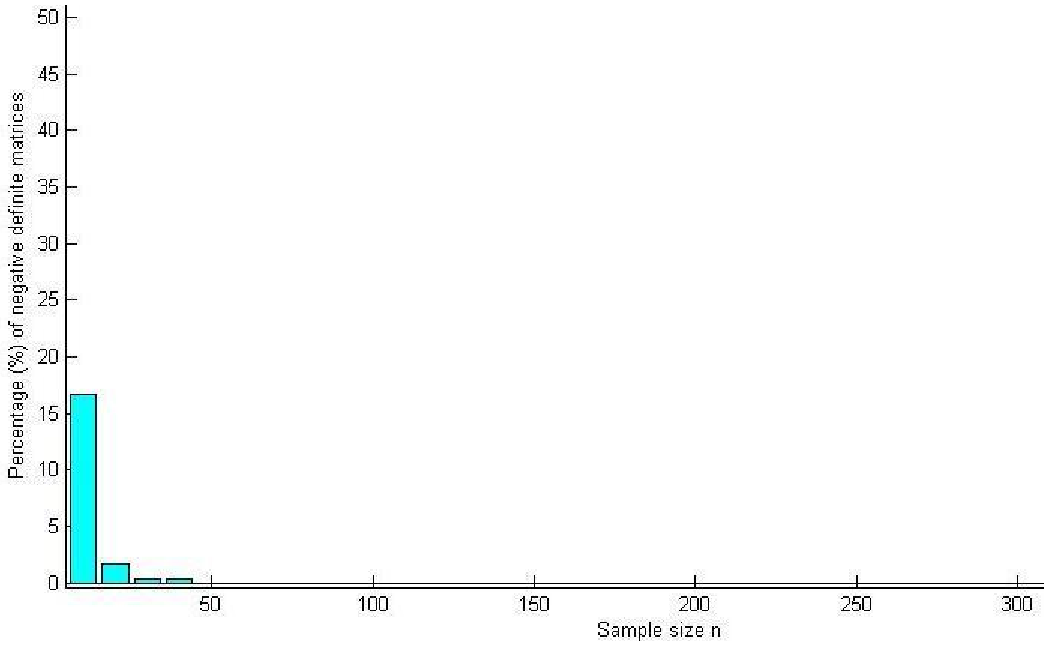

Figure 4.9: Toeplitz covariance structure with different variances

is given by ($n = 10$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 3.9652 & 0.9617 & 1.9537 & 0.9617 \\ 0.9617 & 3.9652 & 0.9617 & 1.9537 \\ 1.9537 & 0.9617 & 3.9652 & 0.9617 \\ 0.9617 & 1.9537 & 0.9617 & 3.9652 \end{pmatrix}$$

and ($n = 100$)

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 3.9996 & 1.0065 & 1.9985 & 1.0065 \\ 1.0065 & 3.9996 & 1.0065 & 1.9985 \\ 1.9985 & 1.0065 & 3.9996 & 1.0065 \\ 1.0065 & 1.9985 & 1.0065 & 3.9996 \end{pmatrix}$$

In this case, the estimates of $\mathbf{\Sigma}$ are closely to the true value.
The circular toeplitz covariance structure with $p = 4$ shows a zero percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for all sample size n considered (see table 4.2).

Circular toeplitz covariance structure (with $p = 5$)

When the circular toeplitz covariance structure (with $p = 5$) is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 4 & 1 & 2 & 2 & 1 \\ 1 & 4 & 1 & 2 & 2 \\ 2 & 1 & 4 & 1 & 2 \\ 2 & 2 & 1 & 4 & 1 \\ 1 & 2 & 2 & 1 & 4 \end{pmatrix}$$

is given by $(n = 10)$

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 3.9928 & 1.0134 & 2.0089 & 2.0089 & 1.0134 \\ 1.0134 & 3.9928 & 1.0134 & 2.0089 & 2.0089 \\ 2.0089 & 1.0134 & 3.9928 & 1.0134 & 2.0089 \\ 2.0089 & 2.0089 & 1.0134 & 3.9928 & 1.0134 \\ 1.0134 & 2.0089 & 2.0089 & 1.0134 & 3.9928 \end{pmatrix}$$

and $(n = 500)$

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 3.9951 & 0.9956 & 1.9973 & 1.9973 & 0.9956 \\ 0.9956 & 3.9951 & 0.9956 & 1.9973 & 1.9973 \\ 1.9973 & 0.9956 & 3.9951 & 0.9956 & 1.9973 \\ 1.9973 & 1.9973 & 0.9956 & 3.9951 & 0.9956 \\ 0.9956 & 1.9973 & 1.9973 & 0.9956 & 3.9951 \end{pmatrix}$$

In this case, the estimates of $\mathbf{\Sigma}$ are closely to the true value.
The circular toeplitz covariance structure with $p = 5$ also shows a zero percentage of non positive definite matrices of the estimate of the covariance matrix $\mathbf{\Sigma}$ for all sample size n (see table 4.2).

Intraclass covariance structure

When the intraclass covariance structure is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

is given by $(n = 10)$

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} 1.9930 & 0.9939 & 0.9939 & 0.9939 \\ 0.9939 & 1.9930 & 0.9939 & 0.9939 \\ 0.9939 & 0.9939 & 1.9930 & 0.9939 \\ 0.9939 & 0.9939 & 0.9939 & 1.9930 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 2.0028 & 1.0008 & 1.0008 & 1.0008 \\ 1.0008 & 2.0028 & 1.0008 & 1.0008 \\ 1.0008 & 1.0008 & 2.0028 & 1.0008 \\ 1.0008 & 1.0008 & 1.0008 & 2.0028 \end{pmatrix}$$

In this case, the estimates of $\boldsymbol{\Sigma}$ are closely to the true value.
The intraclass covariance structure also shows zero percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for all sample size n (see table 4.2).

Compound symmetric type I structure

When the compound symmetric type I structure is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 10 & 9 & 9 \\ 1 & 9 & 10 & 9 \\ 1 & 9 & 9 & 10 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 2.3202 & 0.7483 & 0.7483 & 0.7483 \\ 0.7483 & 13.5256 & 12.4115 & 12.4115 \\ 0.7483 & 12.4115 & 13.5256 & 12.4115 \\ 0.7483 & 12.4115 & 12.4115 & 13.5256 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 1.9999 & 1.0000 & 1.0000 & 1.0000 \\ 1.0000 & 10.0141 & 9.0141 & 9.0141 \\ 1.0000 & 9.0141 & 10.0141 & 9.0141 \\ 1.0000 & 9.0141 & 9.0141 & 10.0141 \end{pmatrix}$$

In this case, the estimates of $\boldsymbol{\Sigma}$ are closely to the true value.
The compound symmetric type I structure shows a small percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for sample size around $n = 10$ (see table 4.2).

Compound symmetric type II structure

When the compound symmetric type II structure is considered, the estimate of the positive definite covariance matrix

$$\Sigma = \begin{pmatrix} 5 & 2 & 3 & 1 \\ 2 & 5 & 1 & 3 \\ 3 & 1 & 6 & 4 \\ 1 & 3 & 4 & 6 \end{pmatrix}$$

is given by ($n = 10$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 37.0450 & -3.9938 & -26.6162 & 19.2260 \\ -3.9938 & 37.0450 & 19.2260 & -26.6162 \\ -26.6162 & 19.2260 & 63.9544 & -1.1993 \\ 19.2260 & -26.6162 & -1.1993 & 63.9544 \end{pmatrix}$$

and ($n = 500$)

$$\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 5.0083 & 2.0098 & 2.9944 & 0.9961 \\ 2.0098 & 5.0083 & 0.9961 & 2.9944 \\ 2.9944 & 0.9961 & 5.9886 & 3.9884 \\ 0.9961 & 2.9944 & 3.9884 & 5.9886 \end{pmatrix}$$

For the above two estimates of $\boldsymbol{\Sigma}$, for small sample size $n = 10$ the estimate is not closed to the proposed value of $\boldsymbol{\Sigma}$.
The compound symmetric type II structure shows a small percentage of non positive definite matrices of the estimate of the covariance matrix $\boldsymbol{\Sigma}$ for small sample size n around $n = 20$ (see table 4.2).

The table (4.2) shows the percentage of non positive definite estimates of the linearly structured covariance matrix $\mathbf{\Sigma}$ for the different sample sizes and linear structures for the extended growth curve model (EGCM);

Table 4.2: Percentage of non positive definite estimates of $\mathbf{\Sigma}$ for the different sample sizes and linear structures for EGCM

Sample size	Percentage of non positive definite estimates of $\mathbf{\Sigma}$								
n	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**
10	29.7	42.6	33.2	16.7	0	0	0	2.9	5.0
20	3.7	11.3	21.6	1.7	0	0	0	0	0
30	0.7	4.4	15.1	0.3	0	0	0	0	0
40	0.2	1.2	11.5	0.3	0	0	0	0	0
50	0	1.1	8.2	0	0	0	0	0	0
60	0.1	0.2	5.5	0	0	0	0	0	0
70	0	0	3.5	0	0	0	0	0	0
80	0	0	3.6	0	0	0	0	0	0
90	0	0	1.9	0	0	0	0	0	0
100	0	0	1.4	0	0	0	0	0	0
110	0	0	1.4	0	0	0	0	0	0
120	0	0	1.2	0	0	0	0	0	0
130	0	0	1.3	0	0	0	0	0	0
140	0	0	0.2	0	0	0	0	0	0
150	0	0	1.1	0	0	0	0	0	0
160	0	0	0.2	0	0	0	0	0	0
170	0	0	0	0	0	0	0	0	0
180	0	0	0.4	0	0	0	0	0	0
190	0	0	0	0	0	0	0	0	0
200	0	0	0.4	0	0	0	0	0	0
210	0	0	0	0	0	0	0	0	0
220	0	0	0	0	0	0	0	0	0
230	0	0	0.1	0	0	0	0	0	0
240	0	0	0	0	0	0	0	0	0
250	0	0	0	0	0	0	0	0	0
260	0	0	0.1	0	0	0	0	0	0
270	0	0	0.1	0	0	0	0	0	0
280	0	0	0	0	0	0	0	0	0
290	0	0	0	0	0	0	0	0	0
300	0	0	0.1	0	0	0	0	0	0

Chapter 5

Conclusions and Suggestions for the further works

In this master thesis entitled ***Estimation of parameters in the growth curve model with a linearly structured covariance matrix− A simulation study***, we have tried to identify which kind of structures, the algorithm proposed by [5] and [6] produce positive definite estimates for the linearly structured covariance matrix $\mathbf{\Sigma}$. In this master thesis, it was found that, those algorithms produce consistent estimator of the linearly structured covariance matrix and perform well in general for large samples for all classes of linear structures studied since they show 100% of positive definite estimates of the covariance matrix. The class of circular toeplitz covariance structures, and intraclass covariance structure show 100% of positive definite estimates of the covariance matrix $\mathbf{\Sigma}$ for both simple and extended growth curve models for all sample sizes, but compound symmetry type I and II structure show a small percentage (around zero) of non positive definite estimates for sample sizes $n = 10$. Also, the banded structures and toeplitz covariance structure (with the same variances) show a non zero percentage of non positive definite estimates for the covariance matrix $\mathbf{\Sigma}$ for small and large sample sizes.

The other important observation is that, for the extended growth curve model, the percentages of non positive definite estimates of the covariance matrix $\mathbf{\Sigma}$ are higher than the ones for the simple growth curve model for the linear structures studied except for compound symmetry ($typeI\&II$).

The covariance structures are classified according to the graph below

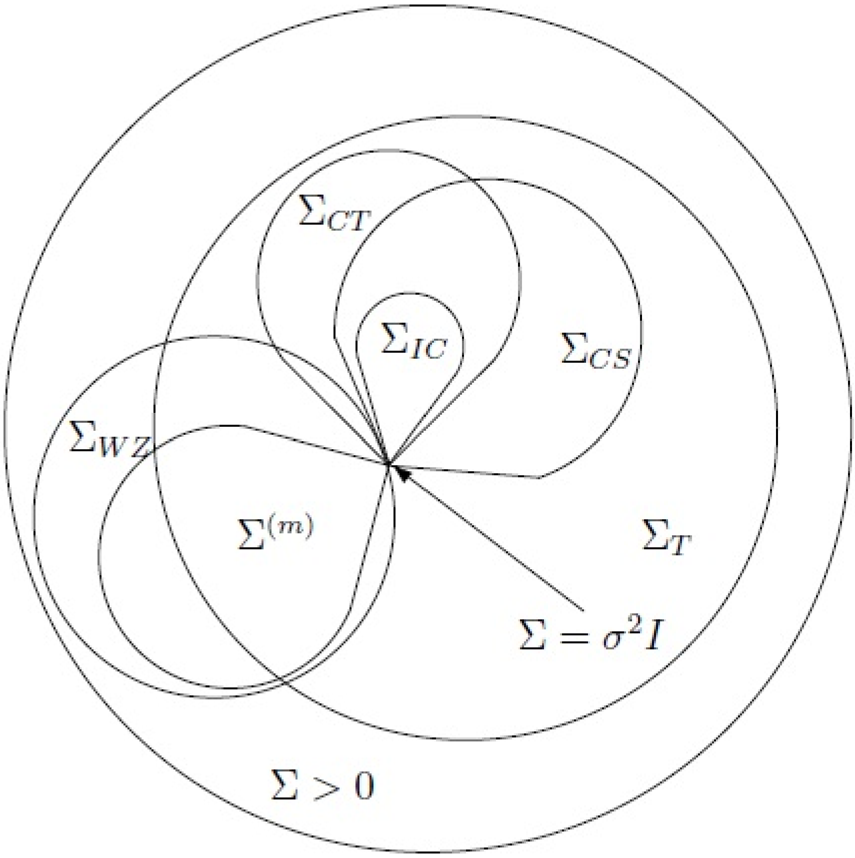

Figure 5.1: Different covariance structures. (Σ_{WZ} = *with zeros*, $\Sigma^{(m)}$ = *banded*, Σ_T = *toeplitz*, Σ_{CT} = *circular toeplitz*, Σ_{CS} = *compound symmetry* and Σ_{IC} = *intraclass*)

At the completion of this thesis some points have to be pointed out as suggestion for the further works:

- For simulation study, if there is a need of the use of a linear structure, it is recommended to use the circular toeplitz covariance structure, or intraclass covariance structure where it is clear that always the estimates of the covariance matrix $\mathbf{\Sigma}$ are positive definite.

- In further research, it is also very interesting to propose other algorithms that can always produce the positive definite estimate of the covariance matrix $\mathbf{\Sigma}$ for all structures for large sample size as well as for small sample size.
- Understand why some linear structures, for example circular toeplitz which also include intraclass, always give positive definite estimates of the covariance matrix $\mathbf{\Sigma}$ and other not.

Bibliography

[1] Chaudhuri, S., Drton, M., and Richardson, T. S. (2007).*Estimation of a covariance matrix with zeros.* Biometrika, 94(1) : $199 - 216$.

[2] Dutilleul, P. (1999). *The MLE algorithm for the matrix normal distribution.* Journal of Statistical Computation and Simulation, 64(2) : $105 - 123$.

[3] Khatri, C. G. (1966*a*). *A note on a MANOVA model applied to problems in growth curve.* Annals of the Institute of Statistical Mathematics, 18(1) : $75 - 86$.

[4] Kollo,T. and von Rosen, D. (2005). *Advanced Multivariate Statistics with matrices.* Springer, New York.

[5] Nzabanita, J., von Rosen, D., and Singull, M. (2012).*Estimation of parameters in the extended growth curve model with a linearly structured covariance matrix.* Acta et Commentationes Universitatis Tartuensis de Mathematica, 16(1) : $13 - 32$.

[6] Ohlson, M. and von Rosen, D. (2010). *Explicit estimators of parameters in the growth curve model with linearly structured covariance matrices.* Journal of Multivariate Analysis, 101 : $1284 - 1295$.

[7] Ohlson, M., Ahmad, M. R., and von Rosen, D. (2011). *The multilinear normal distribution: Introduction and some basic properties.* Journal of Multivariate Analysis, 113 : $37 - 47$.

[8] Olkin, I. and Press, S. (1969). *Testing and estimation for a circular stationary model.* The Annals of Mathematical Statistics, 40(4) : $1358 - 1373$.

[9] Pan, Jian-Xin and Fang, Kai-Tai (2002). *Growth curve models and statistical diagnostics.* Springer Series in Statistics. New York: Springer-Verlag. ISBN $0 - 387 - 95053 - 2$.

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

[11] Seber, G. A. F. and Wild, C. J. (1989). *Growth models. Nonlinear regression*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons, Inc. pp. 325 − 367. ISBN 0 − 471 − 61760 − 1.

[12] Seid Hamid, J. and von Rosen, D. (2006). *Residuals in the extended growth curve model*. Scandinavian Journal of Statistics, 33(1) : 121−138.

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

[14] Srivastava, M.S. and von Rosen, D. (1999). *Growth curve model*. In: Multivariate Analysis, Design of Experiments, and Survey Sampling. (S. Ghosh, ed.), Statistics: Textbooks and Monographs 159, Dekker, New York, pp. 547 − 578.

[15] Srivastava, M. S. and von Rosen, D. (2008). *Models with a Kronecker product covariance structure: estimation and testing*. Mathematical Methods in Statistics, 17 : 357 − 370.

[16] Szatrowski, T. H. (1982). *Testing and estimation in the block compound symmetry problem*. Journal of Educational Statistics, 7(1) : 3 − 18.

[17] Verbyla, A. and Venables, W. (1988). *An extension of the growth curve model*. Biometrica.

[18] von Rosen, D. (1989). *Maximum likelihood estimators in multivariate linear normal models*. Journal of Multivariate Analysis, 31 : 187 − 200.

[19] von Rosen, D., (1995), *Residuals in the growth curve model*. Annals of the Institute of Statistical Mathematics, 47(1) : 129 − 136.

[20] Votaw, D. F. (1948). *Testing compound symmetry in a normal multivariate distribution*. The Annals of Mathematical Statistics, 19(4) : 447 − 473.

[21] Wilks, S. S. (1946). *Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution*. The Annals of Mathematical Statistics, 17(3) : 257 − 281.

[22] Yokoyama, T. (1996). *Extended Growth Curve models with random-effects covariance structures*. Communications in Statistics - Theory and Methods, 25(3) : 571 − 584.