Contributions to Estimation and Testing Block Covariance Structures in Multivariate Normal Models

Yuli Liang

Stockholm
University

Contributions to Estimation and Testing Block Covariance Structures in Multivariate Normal Models

Yuli Liang

Abstract

This thesis concerns inference problems in balanced random effects models with a so-called block circular Toeplitz covariance structure. This class of covariance structures describes the dependency of some specific multivariate two-level data when both compound symmetry and circular symmetry appear simultaneously.

We derive two covariance structures under two different invariance restrictions. The obtained covariance structures reflect both circularity and exchangeability present in the data. In particular, estimation in the balanced random effects with block circular covariance matrices is considered. The spectral properties of such patterned covariance matrices are provided. Maximum likelihood estimation is performed through the spectral decomposition of the patterned covariance matrices. Existence of the explicit maximum likelihood estimators is discussed and sufficient conditions for obtaining explicit and unique estimators for the variance-covariance components are derived. Different restricted models are discussed and the corresponding maximum likelihood estimators are presented.

This thesis also deals with hypothesis testing of block covariance structures, especially block circular Toeplitz covariance matrices. We consider both so-called external tests and internal tests. In the external tests, various hypotheses about testing block covariance structures, as well as mean structures, are considered, and the internal tests are concerned with testing specific covariance parameters given the block circular Toeplitz structure. Likelihood ratio tests are constructed, and the null distributions of the corresponding test statistics are derived.

Keywords: Block circular symmetry, covariance parameters, explicit maximum likelihood estimator, likelihood ratio test, restricted model, Toeplitz matrix

ISBN 978-91-7649-136-2

Printer: Holmbergs, Malmö 2015
Distributor: Department of Statistics, Stockholm University

To my dear family and friends

List of Papers

The thesis includes the following four papers, referred to in the text by their Roman numerals.

PAPER I: Liang, Y., von Rosen, T., and von Rosen, D. (2011). Block circular symmetry in multilevel models. *Research Report* 2011:3, *Department of Statistics, Stockholm University, revised version.*

PAPER II: Liang, Y., von Rosen, D., and von Rosen, T. (2012). On estimation in multilevel models with block circular symmetric covariance structure. *Acta et Commentationes Universitatis de Mathematica*, **16**, 83-96.

PAPER III: Liang, Y., von Rosen, D., and von Rosen, T. (2014). On estimation in hierarchical models with block circular covariance structures. *Annals of the Institute of Statistical Mathematics.* DOI: 10.1007/s10463-014-0475-8.

PAPER IV: Liang, Y., von Rosen, D., and von Rosen, T. (2015). Testing in multivariate normal models with block circular covariance structures. *Research Report* 2015:2, *Department of Statistics, Stockholm University.*

Contents

Acknowledgements

My journey as a doctoral student is now approaching the end. The Chinese poet Xu Zhimo has ever said: "*fortune to have, fate to lose.*"[1] I truly believe that I have been fortunate in choosing statistics as my subject, coming to Sweden, pursuing master's and doctoral studies and meeting a lot of kind-hearted people who have given me help and support in one way or another.

First and foremost, I would like to express my deepest gratitude to my amazing supervisors, Tatjana von Rosen and Dietrich von Rosen. The word supervisor in Swedish is "handledare" and you lead me in the right direction like a beacon. To Tatjana, thank you for introducing to me the interesting and important research problems that are treated in this thesis and guiding me since the first day I was your doctoral student. I could never have accomplished this thesis without your support. To Dietrich, thank you for all of your time spent reading and commenting on my draft essays. I have learned from you not only statistics but also an important attitude for being a researcher: *slow down and focus,* which is invaluable to me.

I am also very thankful to my colleagues at the department. Special thanks to Dan Hedlin, Ellinor Fackle-Fornius, Jessica Franzén, Gebrenegus Ghilagaber, Michael Carlson, Hans Nyquist, Daniel Thorburn, Frank Miller and Per-Gösta Andersson, for your friendliness and valuable suggestions for my research and future career. Big thanks to Jenny Leontine Olsson for being so nice and supportive all the time. I wish to thank Håkan Slättman, Richard Hager, Marcus Berg and Per Fallgren for always being friendly and helpful.

I want to thank Fan Yang Wallentin who suggested to me that I pursue PhD studies. Thanks to Adam Taube for bringing me into the world of medical statistics when I was a master's student in Uppsala. Thanks to Kenneth Carling for your supervision when I wrote my D-essay in Borlänge. Thanks to Mattias Villani, Martin Singull and Jolanta Pielaszkiewicz for all your friendliness and encouragement.

During these years I have been visiting some people around the world. Thanks to Professor Júlia Volaufová for taking the time to discuss my re-

[1]This is a free translation. Xu Zhimo (January 15, 1897 – November 19, 1931) was an early 20th-century Chinese poet.

search, giving me a memorable stay in New Orleans and sharing knowledge during your course “Mixed linear models". Thanks to Professor Thomas Mathew for my visit at the Department of Mathematics & Statistics, University of Maryland, Baltimore County. Thanks to Professor Augustyn Markiewicz for organizing the nice workshop on “Planning and analysis of tensor-experiments" in Będlewo, Poland. My thanks also go to Associate Professor Anuradha Roy for our time working together in San Antonio.

I am grateful for the financial support from the Department of Statistics, the travel grants from Stockholm University and the Royal Swedish Academy of Science.

I am deeply grateful to my friends and fellow doctoral students, former and present, at the department. Thank you for all the joyful conversations and for sharing this experience with me. In particular, I would like to thank Chengcheng; you were my fellow doctoral student from the first day at the department. I enjoyed the days we spent together taking courses, traveling to conferences, visiting other researchers, discussing various statistical aspects and even fighting to meet many deadlines. To Feng, you were even my fellow master student from the first day in Sweden. Thank you for your great friendship during these years. To Bergrún, thank you for all the good times in Stockholm especially on training and dinners, which have become my nice memories. To Karin, Olivia, Sofia and Annika, thanks all of you for providing such a pleasant work environment and especially thank you for your support and comforting words when difficulties came.

I would also like to thank other friends who have made my life in Sweden more enjoyable. Ying Pang, thank you for taking care of me and being good company in Stockholm. To Xin Zhao, thank you for the precious friendship you provided since the first day I entered the university. To Ying Li, I still remember the days in China when we studied together and your persistence inspired me a lot. To Hao Luo, Dao Li, Xijia Liu, Jianxin Wei, Xia Shen and Xingwu Zhou, thank you for setting a good example for me concerning life as a PhD student. To Jia, Qun, Yamei and Cecilia, I have enjoyed all the wonderful moments with you. To Xiaolu and Haopeng, thank you for your kindness every time we have met.

Finally, I really appreciate all the love my dear family has given to me. To my parents, it was you who made me realize the power of knowledge. Thank you for always backing me up and giving me the courage to study abroad. My final appreciation goes to my husband, Deliang. Thank you for believing in me, encouraging me and putting a smile on my face every single day.

Yuli Liang
Stockholm, March 2015

1. Introduction

A statistical model can be considered as an approximation of a real life phenomenon using probabilistic concepts. In the general statistical paradigm, one starts with a specification of a relatively simple model that describes reality as close as possible. This may be according to substantive theories or based on a practitioners' best knowledge. The forthcoming issue concerns statistical inference of the specified model, which can be a multivariate type when modeling multiple response variables jointly, e.g. parameter estimation and hypothesis testing.

1.1 Background

In statistics, the concept of covariance matrix, also called dispersion matrix or variance-covariance matrix, plays a crucial role in statistical modelling since it is a tool to describe the underlying dependency between two or more sets of random variables. In this thesis, patterned covariance matrices are studied. Briefly speaking, a patterned covariance matrix means that besides the restrictions, symmetry and positive semidefiniteness, there exist some more restrictions. For example, very often there exists some theoretical justification, which tells us that the assumed covariance structure is not arbitrary but following a distinctive pattern (Fitzmaurice *et al.*, 2004). For example, in certain experimental designs, when the within-subject factor is randomly allocated to subjects, the model assumption may include a covariance matrix, where all responses have the same variance and any pair of responses have the same covariance. This type of covariance matrix is called compound symmetry (CS) structure, which also is called equicorrelation structure, uniformly structure or intraclass structure. In some longitudinal studies, the covariance matrix can assume that any pair of responses that are equally separated in time have the same correlation. This pattern is referred to as a Toeplitz structure. There are some special kinds of Toeplitz matrices that are commonly used in practice. One is the first-order autoregressive structure, abbreviated as AR(1), where the correlations decline over time as the separation between any pairs of observations increases. The other is a banded Toeplitz matrix, also called q-dependent structure, where all covariances more than q steps apart equal zero. A third special case of a Toeplitz matrix is the symmetric circular Toeplitz (CT) matrix, where the correlation between two measurements only depends on their distance, or we may say it depends on the number of observations between them.

Considerable attention has been paid to studies of patterned covariance matrices because comparing with the $\frac{p(p+1)}{2}$ unknown parameters in a $p \times p$ unstructured covariance matrix, many covariance structures are fairly parsimonious. It follows that both CS and AR(1) covariance structures only have 2 unknown parameters, while the Toeplitz matrix has p parameters, the banded Toeplitz matrix has q parameters ($q < p$) and the symmetric circular Toeplitz matrix has $\left[\frac{p}{2}\right] + 1$ parameters, where $[\bullet]$ denotes the integer function. In models including repeated measurements, the number of unknown parameters in the covariance matrix increases rapidly when the number of repeated measurements is increasing. The parsimony is important for a statistical inference, especially when the sample size is small.

The study of the multivariate normal models with patterned covariance matrices can be traced back to Wilks (1946), in connection with some educational problems, and was extended by Votaw (1948) when considering medical problems. Geisser (1963) considered multivariate analysis of variance (MANOVA) for a CS structure and tested the mean vector. Fleiss (1966) studied a "block version" of the CS structure (see 2.9 in Chapter 2) involving a test of reliability. In the 1970's, this area was intensively developed by Olkin (1973b,a), Khatri (1973), Anderson (1973), Arnold (1973) and Krishnaiah and Lee (1974), among others. Olkin (1973b) considered a multivariate normal model with a block circular structure (see (2.12) in Chapter 2) which the covariance matrix exhibits as circularity in blocks. Olkin (1973a) gave a generalized form of the problem considered by Wilks (1946), which stemmed from a problem in biometry. Khatri (1973) investigated the testing problems of certain covariance structures under a growth curve model. Anderson (1973) dealt with multivariate observations where covariance matrix is a linear combination of known symmetric matrices (see (2.1) in Chapter 2). Arnold (1973) studied certain patterned covariance matrices under both the null and alternative hypotheses which can be transformed to "products" of problems where the covariance matrices are not assumed to be patterned. Krishnaiah and Lee (1974) considered the problems of testing hypotheses when the covariance structure follows certain patterns, and one of the hypotheses considered by Krishnaiah and Lee (1974) contains, among others, both block CS structure and block CT structure as special cases.

Although multivariate normal models with patterned covariance matrices were studied entensively many decades ago, there is a variety of questions still to be addressed, due to interesting and challenging problems arising in various applications such as medical and educational studies. Viana and Olkin (2000) considered a statistical model that can be used in medical studies of paired organs. The data came from visual assessments on N subjects at k time points, and the model assumed a correlation between fellow observations. Let y_{t1} and y_{t2} be the observation of the right and left eyes from one person, respectively, at any time points t and u which are vision-symmetric, $t, u = 1, \ldots, k$. Here "symmetry" means that the left-right labeling is irrelevant at each time point, i.e., $\mathrm{Cov}(y_{t1}, y_{u2}) = \mathrm{Cov}(y_{t2}, y_{u1})$. The covariance structure will exhibit a block pattern corresponding to time points with different CS blocks inside.

Nowadays, it is very common to collect data hierarchically. In particular, for each subject, there may be p variables measured at different sites/positions resulting in doubly multivariate data, i.e., multivariate in two levels (Arnold, 1979; Roy and Fonseca, 2012). The variables may have variations that differ within sites/positions and across dependent subjects. In some clinical trial studies for each subject, the measurements can be collected on more than one variable at different body positions repeatedly over time, resulting in triply multivariate data, i.e., multivariate in three levels (Roy and Leiva, 2008). Similar to the two-level case, in three-level multivariate data the variables may have different variations within sites and across both subjects and times, which should be taken into account. This implies the presence of different block structures in the covariance matrices and the inference should take care concerning this.

Now, a balanced random effects model under a normality assumption, which has been studied intensively in this thesis, will be introduced. The model is assumed to have a general mean and a specific covariance structure of the pattern for which derivation will be motivated in Chapter 5, Theorem 5.1.1. Let y_{ijk} be the response from the kth individual at the jth level of the random factor $\boldsymbol{\gamma}_2$ within the ith level of the random factor $\boldsymbol{\gamma}_1$, $i = 1, \ldots, n_2$, $j = 1, \ldots, n_1$ and $k = 1, \ldots, n$. The model is represented by

$$y_{ijk} = \mu + \gamma_{1,i} + \gamma_{2,ij} + \epsilon_{ijk}, \tag{1.1}$$

where μ is the general mean, $\gamma_{1,i}$ is the random effect, $\gamma_{2,ij}$ is the random effect which is nested within $\gamma_{1,i}$ and ϵ_{ijk} is the random error. A balanced case of model (1.1) means that the range of any subscript of the response $\mathbf{y}_k = (y_{ij})$ does not depend on the values of the other subscripts of $\mathbf{y}_k$.

Let $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be a independent random sample from $N_p(\mathbf{1}_p\mu, \boldsymbol{\Sigma})$, where $p = n_2 n_1$. Put $\boldsymbol{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)$. Then, model (1.1) can be written as $\boldsymbol{Y} \sim N_{p,n}(\mu\mathbf{1}_p\mathbf{1}_n', \boldsymbol{\Sigma}, \boldsymbol{I}_n)$, where

$$\mathbf{y}_k = \mathbf{1}_p\mu + (\boldsymbol{I}_{n_2} \otimes \mathbf{1}_{n_1})\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\epsilon}, \quad k = 1, \ldots, n, \tag{1.2}$$

where $\mathbf{y}_i$ is a $p \times 1$ response vector and $\mathbf{1}_{n_1}$ is the column vector of size n_1, having all elements equal to 1. Here, $\boldsymbol{\gamma}_1 \sim N_{n_2}(\mathbf{0}, \boldsymbol{\Sigma}_1)$, $\boldsymbol{\gamma}_2 \sim N_p(\mathbf{0}, \boldsymbol{\Sigma}_2)$ and $\boldsymbol{\epsilon} \sim N_p(\mathbf{0}, \sigma^2 \boldsymbol{I}_p)$ are assumed to be mutually independent. Furthermore, we assume that both $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are positive semidefinite. Denote $\boldsymbol{Z}_1 = \boldsymbol{I}_{n_2} \otimes \mathbf{1}_{n_1}$. The covariance matrix of $\mathbf{y}_k$ in (1.2) is $\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma} = \boldsymbol{Z}_1 \boldsymbol{\Sigma}_1 \boldsymbol{Z}_1' + \boldsymbol{\Sigma}_2 + \sigma^2 \boldsymbol{I}_p$. In many applications, such as clinical studies, it is crucial to take into account the variations due to the random factor of $\boldsymbol{\gamma}_2$ (e.g., sites/positions) and across the random factor $\boldsymbol{\gamma}_1$ (e.g., time points), in addition to the variations of $\boldsymbol{\gamma}_1$ itself. Moreover, the dependency that nestedness creates may cause different patterns in the covariance matrix, which can be connected to one or several hierarchies or levels.

The covariance matrix of $\mathbf{y}_k$ in (1.2), i.e., $\boldsymbol{\Sigma}$, may have different structures depending on $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. In this thesis, we assume that the covariance $\boldsymbol{\Sigma}$ from model (1.2), equals

$$\boldsymbol{\Sigma} = \boldsymbol{Z}_1 \boldsymbol{\Sigma}_1 \boldsymbol{Z}_1' + \boldsymbol{\Sigma}_2 + \sigma^2 \boldsymbol{I}_p, \tag{1.3}$$

where

$$\boldsymbol{\Sigma}_1 = \sigma_1 \boldsymbol{I}_{n_2} + \sigma_2 (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}), \tag{1.4}$$

$$\boldsymbol{\Sigma}_2 = \boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}, \tag{1.5}$$

$\boldsymbol{J}_{n_2} = \mathbf{1}_{n_2} \mathbf{1}_{n_2}'$ and $\boldsymbol{\Sigma}^{(h)}$ is a CT matrix, $h = 1, 2$, (see also Paper II, Equation (2.5), p.85 or Paper III, p.3). Furthermore, it can be noticed that $\boldsymbol{\Sigma}$ has the same structure as $\boldsymbol{\Sigma}_2$ but with more parameters involved. It is worth observing that model (1.2) is overparametrized, and hence the estimation of parameters in $\boldsymbol{\Sigma}$ faces the problem of identifiability. A parametric statistical model is said to be identified if there is one and only one set of parameters that produces a given probability distribution for the observed variables. Identifiability of model (1.2) will be one of the main concerns in this thesis (see Paper III).

The usefulness of the covariance structure given in (1.3) can appear when modelling phenomena in physical, medical and psychological contexts. Next, we provide some examples arising from different applications that illustrate potential utilization of the model (1.2).

Example 1 Olkin and Press (1969) studied a physical problem concerning modelling signal strength. A point source with a certain number vertices from which a signal received from a satellite is transmitted. Assuming that the signal strength is the same in all directions along the vertices, and the correlations only depend on the number of vertices in between (see Figure 1.1), one would expect a CT structure for underlying dependency between the messages received by the receivers placed at these vertices. Moreover, those messages could be recorded from a couple of exchangeable geocenters which are random samples from a region so that the data can have the circulant property in the receiver (vertices) dimension and a symmetric pattern in the geocenter dimension.

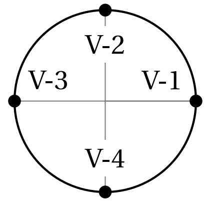

Figure 1.1: A circular structure of the signal receiver with 4 vertices: V-i represents the ith vertex, $i = 1, \ldots, 4$.

Example 2 Louden and Roy (2010) gave one example of the use of the circular symmetry model, which aimed to facilitate the classification of patients suffering in particular from Alzheimer's disease using positron emission tomography (PET) imaging. A healthy brain shows normal metabolism levels throughout the scan, whereas low metabolism in the temporal and parietal lobes on both sides of the brain is seen in patients with Alzheimer's disease. In their study, the three measurements have been taken from temporal lobes, i.e. the anterior temporal, mid temporal and post temporal regions of each temporal lobe. Viewed from the top of the head these three regions in the two hemispheres of the brain seem to form a circle inside the skull, and Louden and Roy (2010) suggested that these six measurements have a CT covariance matrix. The response consists of six measurements (metabolism levels) from the ith patient within kth municipality. Assuming that those patients who received PET imaging are exchangeable and the municipalities are independent samples, the covariance structure can be assumed to have the pattern in (1.3). Although, PET imaging from different patients are independent of each other, i.e., $\mathbf{\Sigma}^{(2)}$ in $\mathbf{\Sigma}$ is zero matrix.

Example 3 The theory of human values proposed by Schwartz (Schwartz, 1992) is that the ten proposed values, i.e., achievement, hedonism, stimulation, self-direction, universalism, benevolence, tradition, conformity, security, and power, form a circular structure (see Davidov and Depner, 2011, Figure 1), in which values expressing similar motivational goals are close to each other and move farther apart as their goals diverge (Steinmetz *et al.*, 2012). Similarly, there exists a "circle reasoning" when studying interpersonal psychology, e.g., classifying persons into typological categories defined by the coordinates of the interpersonal circle (see Gurtman, 2010, Figure 18.2). Those substantive theories result in, when some assessments are conducted from the sampling subjects, the collected measurements, e.g., the scores of the ten values for an individual; these will be circularly correlated within subjects and equicorrelated between subjects.

1.2 Aims of the thesis

The general purpose of this thesis is to study the problems of estimation and hypothesis testing in multivariate normal models related to the specific block covariance structure $\boldsymbol{\Sigma}$ in model (1.2), namely a block circular Toeplitz structure, which can be used to characterize the dependency of some specific two-level multivariate data.

The following specific aims have been in focus.

- The first aim is to derive a block covariance structure which can model the dependency of a specific symmetric two-level multivariate data. Here the concept of symmetry or, in other words, invariance, means that the covariance matrix will remain unchanged (invariant) under certain orthogonal transformations (e.g. permutation).
- The second aim is to obtain estimators for the parameters of model (1.2) with the block circular Toeplitz covariance structure given in (1.3). The focus is on deriving explicit maximum likelihood estimators.
- The third aim is to develop tests for testing different types of symmetry in the covariance matrix as well as testing the mean structure.
- The fourth aim is to construct tests for testing hypotheses about specific parameters in the block circular Toeplitz covariance structure.

1.3 Outline of the thesis

This thesis is organized as follows. In Chapter 1, a general introduction and background of the topic considered in the thesis are given. Chapter 2 focuses on various patterned covariance matrices, especially block covariance structures, which are of primary interest in this thesis. The concept of the symmetry (invariance) model with some simple examples are presented. Chapter 3 provides some existing results on the explicit MLEs for both mean and (co)variance parameters in a multivariate normal model setting. Furthermore, spectral properties of the covariance structures are studied here since they play crucial roles for statistical inference in these models. Chapter 4 provides existing results of the likelihood ratio test (LRT) procedure on some block covariance structures as well as the approximation of the null distributions of the corresponding test statistics. Then some existing methods of testing variance parameters will also be introduced. Summaries of the four papers are given in Chapter 5 where the main results of this thesis will be highlighted. Concluding remarks together with some future research problems appear in the last chapter.

2. Patterned covariance matrices

This chapter is devoted to a brief presentation of the patterned covariance matrices used in statistical modelling. We start with an introduction of both linear and non-linear covariance structures.

2.1 Linear and non-linear covariance structures

According to Anderson (1973), a linear covariance structure is a structure such that the covariance matrix $\boldsymbol{\Sigma} : p \times p$ can be represented as a linear combination of known symmetric matrices:

$$\boldsymbol{\Sigma} = \sum_{i=1}^{s} \sigma_i \boldsymbol{G}_i, \tag{2.1}$$

where $\boldsymbol{G}_1, \ldots, \boldsymbol{G}_s$ are linearly independent, known symmetric matrices and the coefficients σ_i are unknown parameters. Moreover, there is at least one set $\sigma_1, \ldots, \sigma_s$ such that (2.1) is positive definite. The linear independence of $\boldsymbol{G}_i$ leads to all unknown parameters being identifiable means that they can be estimated uniquely.

The concept of linear covariance structure will now be illustrated with the following examples. Recall the various covariance matrices introduced in Chapter 1. The CS structure has the form

$$\boldsymbol{\Sigma}_{CS} = \begin{pmatrix} a & b & \cdots & b \\ b & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots & b & a \end{pmatrix},$$

where a is the variance, b is the covariance and $\boldsymbol{\Sigma}$ is nonnegative definite if and only if $a \geq b \geq -\frac{1}{p-1}a$. The CS structure can be written as

$$\boldsymbol{\Sigma}_{CS} = a\boldsymbol{I}_p + b(\boldsymbol{J}_p - \boldsymbol{I}_p) = \left[a + (p-1)b\right]\boldsymbol{P}_{\mathbf{1}_p} + (a-b)(\boldsymbol{I}_p - \boldsymbol{P}_{\mathbf{1}_p}), \tag{2.2}$$

where $\boldsymbol{P}_{\mathbf{1}_p}$ is the orthogonal projection onto the column space of $\mathbf{1}_p$. Expression (2.2) shows that the CS structure is a linear covariance structure.

The Toeplitz structure is of the form

$$\boldsymbol{\Sigma}_{Toep} = \begin{pmatrix} t_0 & t_1 & t_2 & \cdots & t_{p-1} \\ t_1 & t_0 & t_1 & \cdots & t_{p-2} \\ t_2 & t_1 & t_0 & \cdots & t_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{p-1} & t_{p-2} & t_{p-3} & \cdots & t_0 \end{pmatrix},$$

where t_0 is the variance for all observations and the covariance between any pair of observations (i,j) equals t_{i-j}. Next, let us define a so-called symmetric Toeplitz matrix $ST(p,k)$ in the following way:

$$ST(p,k) = Toep(\overbrace{\underbrace{0,\ldots,0}_{k},1,0,\ldots,0}^{p}),$$

or equivalently

$$(ST(p,k))_{ij} = \begin{cases} 1, & \text{if } |i-j| = k, \\ 0, & \text{otherwise,} \end{cases}$$

where $k \in \{1,\ldots,p-1\}$. For notational convenience denote $ST(p,0) = \boldsymbol{I}_p$. The Toeplitz structure can then be expressed as

$$\boldsymbol{\Sigma}_{Toep} = \sum_{k=0}^{p-1} t_k ST(p,k),$$

and $ST(p,k)$ are linearly independent, $k = 1,\ldots,p-1$. Therefore, the Toeplitz structure is a linear structured covariance matrix and it is also called a linear Toeplitz structure (Marin and Dhorne, 2002). As one of the special cases of the Toeplitz structure, the CT structure can be expressed as

$$\boldsymbol{\Sigma}_{CT} = \begin{pmatrix} t_0 & t_1 & t_2 & \cdots & t_1 \\ t_1 & t_0 & t_1 & \cdots & t_2 \\ t_2 & t_1 & t_0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & \cdots & t_1 & t_0 \end{pmatrix}, \tag{2.3}$$

where t_0 is the variance for all observations and the covariance between any pair of observations (i,j) equals $t_{min\{i-j,n-(i-j)\}}$. The CT structure can be expressed as

$$\boldsymbol{\Sigma}_{CT} = \sum_{k=0}^{[p/2]} t_k SC(p,k), \tag{2.4}$$

where $SC(p,k)$ is called a symmetric circular matrix and is defined as follows:

$$SC(p,k) = Toep(\overbrace{\underbrace{0,\ldots,0}_{k},1,0,\ldots,0,1,\underbrace{0,\ldots,0}_{k-1}}^{p}) \tag{2.5}$$

or equivalently

$$(SC(p,k))_{ij} = \begin{cases} 1, & \text{if } |i-j| = k \text{ or } |i-j| = p-k, \\ 0, & \text{otherwise,} \end{cases}$$

where $k \in \{1,\ldots,[p/2]\}$. For notational convenience denote $SC(p,0) = \boldsymbol{I}_p$.

A non-linear covariance structure basically refers to the non-linear structure of the covariance matrix $\boldsymbol{\Sigma}$ in its parameters. One example is the AR(1) structure:

$$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{p-1} \\ \rho & 1 & \rho & \cdots & \rho^{p-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{p-1} & \rho^{p-2} & \rho^{p-3} & \cdots & 1 \end{pmatrix},$$

where $\rho^k = \mathrm{Cor}(y_j, y_{j+k})$ for all j and k and $\rho \geq 0$.

For some of the above mentioned covariance structures it is not possible to obtain explicit MLEs, for example, the AR(1) and the symmetric Toeplitz covariance matrices. Estimation of both linear and non-linear covariance structures under a normality assumption has been considered by several authors. Ohlson *et al.* (2011) proposed an explicit estimator for an m-dependent covariance structure that is not MLE. The estimator is based on factorizing the full likelihood and maximizing each term separately. For models with a linear Toeplitz covariance structure, Marin and Dhorne (2002) derived a necessary and sufficient condition to obtain an optimal unbiased estimator for any linear combination of the variance components. Their results were obtained by means of commutative Jordan algebras. In Chapter 3, the explicit estimation of patterned covariance matrices will be considered in detail.

2.2 Symmetry models

To have a specific covariance structure in a model means that certain restrictions are imposed on the covariance matrix. In this thesis, we are interested in some specific structures when certain invariance conditions are fulfilled, i.e. when the process generating is supposed to follow a probability distribution whose covariance is invariant with respect to certain orthogonal transformations. Andersson (1975) and Andersson and Madsen (1998) have presented a comprehensive theory of group invariance in multivariate normal models. In the review article of Perlman (1987), the terminology "group symmetry" is used to describe group invariance. The following definition describes the concept of invariance more formally.

Definition 2.2.1 *(Perlman, 1987) Let $\mathcal{G}$ be a finite group of orthogonal transformations. A symmetry model determined by the group $\mathcal{G}$ is a family of models with positive definite covariance matrices*

$$S_{\mathcal{G}} = \{\boldsymbol{\Sigma} | \boldsymbol{G}\boldsymbol{\Sigma}\boldsymbol{G}' = \boldsymbol{\Sigma} \text{ for all } \boldsymbol{G} \in \mathcal{G}\}. \tag{2.6}$$

The covariance matrix $\boldsymbol{\Sigma}$ defined in (2.6) is said to be $\mathcal{G}$-invariant. If $\mathbf{y}$ is a random vector with $Cov(\mathbf{y}) = \boldsymbol{\Sigma}$, then $Cov(\boldsymbol{G}\mathbf{y}) = \boldsymbol{G}\boldsymbol{\Sigma}\boldsymbol{G}'$. Thus, the condition $\boldsymbol{G}\boldsymbol{\Sigma}\boldsymbol{G}' = \boldsymbol{\Sigma}$ in (2.6) implies that $\mathbf{y}$ and $\boldsymbol{G}\mathbf{y}$ have the same covariance matrix. The general theory for symmetry models specified by (2.6) is provided by Andersson (1975). It tells us how $S_{\mathcal{G}}$ should look like, but does not tell us how to derive the particular form of $S_{\mathcal{G}}$ (Eaton, 1983). It is not obvious, given a structure for the covariance matrix, to find the corresponding $\mathcal{G}$, or even to decide whether there is a corresponding $\mathcal{G}$. Nevertheless, given the group, it is possible to find the corresponding $\mathcal{G}$-invariant structure of $\boldsymbol{\Sigma}$ (Marden, 2012). Perlman (1987) discussed and summarized results related to group symmetry models, in which some cases were studied in detail such as spherical symmetry (Mauchly, 1940), complete symmetry (Wilks, 1946), compound symmetry (CS) (Votaw, 1948), circular symmetry (Olkin and Press, 1969), and block circular symmetry (Olkin, 1973b). Moreover, Nahtman (2006), Nahtman and von Rosen (2008) and von Rosen (2011) studied properties of some patterned covariance matrices arising under different symmetry restrictions in balanced mixed linear models.

Our next examples illustrate two symmetry models with different covariance structures: the CS structure and the CT structure given by (2.2) and (2.3), respectively. In order to connect the concept symmetry model with the following examples, we first need to define $\boldsymbol{P}^{(2)}$ to be an $n \times n$ arbitrary permutation matrix, which is an orthogonal matrix whose columns can be obtained by permuting the columns of the identity matrix, e.g.,

$$\boldsymbol{P}^{(2)} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We also define $\boldsymbol{P}^{(1)}$ to be an $n \times n$ arbitrary shift-permutation (SP) matrix (or cyclic permutation matrix) of the form

$$p_{ij}^{(1)} = \begin{cases} 1, & \text{if } j = i + 1 - n\mathbf{1}_{(i>n-1)}, \\ 0, & \text{otherwise,} \end{cases} \tag{2.7}$$

where $\mathbf{1}_{(.)}$ is the indicator function, i.e. $\mathbf{1}_{(a>b)} = 1$ if $a > b$ and $\mathbf{1}_{(a>b)} = 0$ otherwise. For example, when $n = 3$ and $n = 4$, the SP matrices are

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Example 4 Let n measurements be taken under the same experimental conditions, and $\boldsymbol{y} = (y_1, \ldots, y_n)'$ denote the response vector. In some situations, it may be reasonable to suppose that the y_is are exchangeable (with proper assumptions about the mean of $\boldsymbol{y}$). Thus, $(y_1, \ldots, y_n)'$ and $(y_{i_1}, \ldots, y_{i_n})'$, where $(i_1, \ldots, i_n)'$ is any permutation of indices $(1, \ldots, n)$, should have the same covariance structure. Let $\boldsymbol{\Sigma}$ be the covariance matrix of $\boldsymbol{y}$. It has been shown (see Eaton, 1983; Nahtman, 2006) that $\boldsymbol{\Sigma}$ is invariant with respect to all orthogonal transformations defined by $\boldsymbol{P}^{(2)}$ if and only if $\boldsymbol{\Sigma} = (a-b)\boldsymbol{I}_n + b\boldsymbol{J}_n$, where a and b are constants.

Example 5 *(Eaton, 1983)* Consider observations $y_1, \ldots, y_n$, which are taken at n equally spaced points on a circle and are numbered sequentially around the circle. For example, the observations might be temperatures at a fixed cross section on a cylindrical rod when a heat source is present at the center of the rod. It may be reasonable to assume that the covariance between y_j and y_k depends only on how far apart y_j and y_k are on the circle. That is, $\mathrm{Cov}(y_j, y_{j+1})$ does not depend on j, $j = 1, \ldots, n$, where $y_{n+1} \equiv y_1$; $\mathrm{Cov}(y_j, y_{j+2})$ does not depend on j, $j = 1, \ldots, n$, where $y_{n+2} \equiv y_2$; and so on. Assuming that $\mathrm{Var}(y_j)$ does not depend on j, this assumption can be expressed as follows: let $\boldsymbol{y} = (y_1, \ldots, y_n)'$ and $\boldsymbol{\Sigma}$ be the corresponding covariance matrix. Nahtman and von Rosen (2008) have shown that $\boldsymbol{\Sigma}$ is invariant with respect to all orthogonal transformations defined by $\boldsymbol{P}^{(1)}$ in (2.7) if and only if $\boldsymbol{\Sigma}$ is a CT matrix given in (2.3). For example, when $n = 5$, $\boldsymbol{\Sigma}$ is $\boldsymbol{P}^{(1)}$-invariant if and only if

$$\boldsymbol{\Sigma} = \begin{pmatrix} t_0 & t_1 & t_2 & t_2 & t_1 \\ t_1 & t_0 & t_1 & t_2 & t_2 \\ t_2 & t_1 & t_0 & t_1 & t_2 \\ t_2 & t_2 & t_1 & t_0 & t_1 \\ t_1 & t_2 & t_2 & t_1 & t_0 \end{pmatrix}.$$

In the next section, more examples of symmetry models will be given in terms of block structures when certain invariant conditions exist at certain layers of the observations.

2.3 Block covariance structures

The simplest block covariance structure may consist of the following block diagonal pattern:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_0 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Sigma}_0 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\Sigma}_0 \end{pmatrix}, \tag{2.8}$$

where $\boldsymbol{\Sigma}$ is a $up \times up$ matrix and $\boldsymbol{\Sigma}_0$ is an $p \times p$ unstructured covariance matrix for each subject over time. To reduce the number of unknown parameters, especially when p is relatively large, $\boldsymbol{\Sigma}_0$ is usually assumed to have some specific structures, e.g. CS or Toeplitz. The covariance matrix in (2.8) can be considered as a trivial symmetry model, i.e. it is invariant with respect to the identity matrix $\boldsymbol{I}_u \otimes \boldsymbol{I}_p$. The block structure of $\boldsymbol{\Sigma}$ can also be extended to other patterns, for example, the off-diagonal blocks can be included into $\boldsymbol{\Sigma}$ to characterize the dependency between subjects, i.e.,

$$\begin{aligned} \boldsymbol{\Sigma}_{BCS} &= \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \dots & \boldsymbol{\Sigma}_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_0 \end{pmatrix}, \\ &= \boldsymbol{I}_u \otimes \boldsymbol{\Sigma}_0 + (\boldsymbol{J}_u - \boldsymbol{I}_u) \otimes \boldsymbol{\Sigma}_1, \end{aligned} \tag{2.9}$$

where $\boldsymbol{\Sigma}_0$ is a positive definite $p \times p$ covariance matrix and $\boldsymbol{\Sigma}_1$ is a $p \times p$ symmetric matrix, and in order to have $\boldsymbol{\Sigma}_{BCS}$ to be a positive definite matrix, the restriction $\boldsymbol{\Sigma}_0 > \boldsymbol{\Sigma}_1 > -\boldsymbol{\Sigma}_0/(u-1)$ has to be fulfilled (see Lemma 2.1 Roy and Leiva, 2011, for proof), where the notation $\boldsymbol{A} > \boldsymbol{B}$ means that $\boldsymbol{A} - \boldsymbol{B}$ is positive definite. The structure of $\boldsymbol{\Sigma}$ in (2.9) is called block compound symmetry (BCS) and it has been studied by Arnold (1973, 1979) in the general linear model when the error vectors are assumed to be exchangeable and normally distributed. A particular example considered by Olkin (1973a) was the Scholastic Aptitude Tests (SAT) in the USA. Let y_{iV} and y_{iQ} be the score of the verbal part and quantitative part of the SAT test from the i-th year. If the SAT examinations during the successive u years are exchangeable with respect to variations, it implies that

$$\begin{cases} \text{var}(y_{iV}) = \text{var}(y_{iQ}), & \text{for } \forall\ i\text{-th year}, \\ \text{cov}(y_{iV}, y_{jV}) = \text{cov}(y_{iQ}, y_{jQ}), & \text{for } \forall\ i \neq j\text{-th year}, \\ \text{cov}(y_{iV}, y_{jQ}) = \text{cov}(y_{jQ}, y_{iV}), & \text{for } \forall\ i, j\text{-th year}, \end{cases}$$

where $i, j = 1, \ldots, u$. Hence, the joint covariance matrix has the structure given in (2.9).

Recall the concept symmetry model in Section 2.2, it can be shown that $\boldsymbol{\Sigma}_{BCS}$ is invariant with respect to all transformations $\boldsymbol{P}^{(2)} \otimes \boldsymbol{I}_p$, where $\boldsymbol{P}^{(2)}$ is an arbitrary permutation matrix with size $u \times u$.

There is another type of covariance structure which we call double complete symmetric (DCS) structure, i.e.,

$$\boldsymbol{\Sigma}_{DCS} = \boldsymbol{I}_u \otimes \left[a\boldsymbol{I}_p + b(\boldsymbol{J}_p - \boldsymbol{I}_p)\right] + (\boldsymbol{J}_u - \boldsymbol{I}_u) \otimes c\boldsymbol{J}_p. \tag{2.10}$$

One extension of $\boldsymbol{\Sigma}_{DCS}$ is the following block double complete symmetric (BDCS) structure, which is called “jointly equicorrelated covariance" matrix (Roy and Fonseca, 2012):

$$\boldsymbol{\Sigma}_{BDCS} = \boldsymbol{I}_v \otimes \boldsymbol{\Sigma}_{BCS} + (\boldsymbol{J}_v - \boldsymbol{I}_v) \otimes \boldsymbol{J}_u \otimes \boldsymbol{W}, \tag{2.11}$$

where $\boldsymbol{\Sigma}_{BCS}$ is given by (2.9) and $\boldsymbol{W}$ is a $p \times p$ symmetric matrix. In the study of Roy and Fonseca (2012), the matrix $\boldsymbol{\Sigma}_{BDCS}$ is assumed when modelling multivariate three-level data, where $\boldsymbol{\Sigma}_0$ characterizes the dependency of the p responses at any given location and at any given time point and $\boldsymbol{\Sigma}_1$ characterizes the dependency of the p responses between any two locations and at any given time point. The matrix $\boldsymbol{W}$ represents the dependency of the p responses between any two time points and it is the same for any pair of time points. When $v = 2$, we have

$$\boldsymbol{\Sigma}_{BDCS} = \left(\begin{array}{cccc|cccc} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \ldots & \boldsymbol{\Sigma}_1 & \boldsymbol{W} & \boldsymbol{W} & \ldots & \boldsymbol{W} \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \ldots & \boldsymbol{\Sigma}_1 & \boldsymbol{W} & \boldsymbol{W} & \ldots & \boldsymbol{W} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \ldots & \boldsymbol{\Sigma}_0 & \boldsymbol{W} & \boldsymbol{W} & \ldots & \boldsymbol{W} \\ \hline \boldsymbol{W} & \boldsymbol{W} & \ldots & \boldsymbol{W} & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \ldots & \boldsymbol{\Sigma}_1 \\ \boldsymbol{W} & \boldsymbol{W} & \ldots & \boldsymbol{W} & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \ldots & \boldsymbol{\Sigma}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{W} & \boldsymbol{W} & \ldots & \boldsymbol{W} & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_1 & \ldots & \boldsymbol{\Sigma}_0 \end{array}\right)$$

Olkin (1973b) might be the first to discuss circular symmetry in blocks, as an extension of the circularly symmetric model (the CT structure) considered by Olkin and Press (1969). Olkin (1973b) considered the following block circular Toeplitz (BCT) structure:

$$\boldsymbol{\Sigma}_{BCT} = \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \dots & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \dots & \boldsymbol{\Sigma}_3 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \dots & \boldsymbol{\Sigma}_4 & \boldsymbol{\Sigma}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_3 & \boldsymbol{\Sigma}_4 & \dots & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_3 & \dots & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 \end{pmatrix}, \tag{2.12}$$

where every matrix $\boldsymbol{\Sigma}_i$ is a $p \times p$ symmetric matrix, and $\boldsymbol{\Sigma}_0$ is positive definite. It can be shown that $\boldsymbol{\Sigma}_{BCT}$ is invariant with respect to all orthogonal transformations $\boldsymbol{P}^{(1)} \otimes \boldsymbol{I}_p$, where $\boldsymbol{P}^{(1)}$ is the SP matrix given in (2.7). The BCT structure considered in Olkin (1973b) was justified by a physical model in which signals are received at the vertices of a regular polygon. When the signal received at each vertex is characterized by p components, we may have the assumption that the variation coming from each p component depends only on the number of vertices in between. The problem is a “multivariate version" of Example 1 in Chapter 1.

Nahtman (2006) and Nahtman and von Rosen (2008) studied symmetry models arising in K-way tables, which contain k random factors $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_k$, where each factor $\boldsymbol{\gamma}_k$ takes value in a finite set of factor levels. In particular, in the context of a 2-way layout model, Nahtman (2006) studied the covariance structure, with a second-order interaction effect, expressed as

$$\boldsymbol{\Sigma}_{BCS-CS} = \boldsymbol{I}_u \otimes \left[a\boldsymbol{I}_p + b(\boldsymbol{J}_p - \boldsymbol{I}_p)\right] + (\boldsymbol{J}_u - \boldsymbol{I}_u) \otimes \left[c\boldsymbol{I}_p + d(\boldsymbol{J}_p - \boldsymbol{I}_p)\right]. \tag{2.13}$$

Nahtman (2006) has shown that the matrix in (2.13) is invariant with respect to all orthogonal transformations $\boldsymbol{P}_1^{(2)} \otimes \boldsymbol{P}_2^{(2)}$. It is a special case of the BCS structure when both $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ in (2.9) have the CS structures, whereas it has the DCS structure in (2.10) as a special case.

As a follow up study, Nahtman and von Rosen (2008) examined shift permutation in K-way tables. Among others in 2-way tables, it leads to the study of the following block circular Toeplitz matrix with circular Toeplitz blocks inside, denoted as BCT-CT structure:

$$\boldsymbol{\Sigma}_{BCT-CT} = \sum_{k_2=0}^{[u/2]} \sum_{k_1=0}^{[p/2]} t_k SC(u, k_2) \otimes SC(p, k_1), \tag{2.14}$$

where $k = (\frac{p}{2}+1)k_2 + k_1$ and $SC(\bullet, \bullet)$ is the symmetric circular matrix defined by (2.5). For example, when $u = 4$ and $p = 4$, we have

$$
\begin{aligned}
\boldsymbol{\Sigma}_{BCT-CT} &= \begin{pmatrix} \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 & \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_0 \end{pmatrix} \\
&= \left(\begin{array}{cccc|cccc|cccc|cccc}
\tau_0 & \tau_1 & \tau_2 & \tau_1 & \tau_3 & \tau_4 & \tau_5 & \tau_4 & \tau_6 & \tau_7 & \tau_8 & \tau_7 & \tau_3 & \tau_4 & \tau_5 & \tau_4 \\
\tau_1 & \tau_0 & \tau_1 & \tau_2 & \tau_4 & \tau_3 & \tau_4 & \tau_5 & \tau_7 & \tau_6 & \tau_7 & \tau_8 & \tau_4 & \tau_3 & \tau_4 & \tau_5 \\
\tau_2 & \tau_1 & \tau_0 & \tau_1 & \tau_5 & \tau_4 & \tau_3 & \tau_4 & \tau_8 & \tau_7 & \tau_6 & \tau_7 & \tau_5 & \tau_4 & \tau_3 & \tau_4 \\
\tau_1 & \tau_2 & \tau_1 & \tau_0 & \tau_4 & \tau_5 & \tau_4 & \tau_3 & \tau_7 & \tau_8 & \tau_7 & \tau_6 & \tau_4 & \tau_5 & \tau_4 & \tau_3 \\
\hline
\tau_3 & \tau_4 & \tau_5 & \tau_4 & \tau_0 & \tau_1 & \tau_2 & \tau_1 & \tau_3 & \tau_4 & \tau_5 & \tau_4 & \tau_6 & \tau_7 & \tau_8 & \tau_7 \\
\tau_4 & \tau_3 & \tau_4 & \tau_5 & \tau_1 & \tau_0 & \tau_1 & \tau_2 & \tau_4 & \tau_3 & \tau_4 & \tau_5 & \tau_7 & \tau_6 & \tau_7 & \tau_8 \\
\tau_5 & \tau_4 & \tau_3 & \tau_4 & \tau_2 & \tau_1 & \tau_0 & \tau_1 & \tau_5 & \tau_4 & \tau_3 & \tau_4 & \tau_8 & \tau_7 & \tau_6 & \tau_7 \\
\tau_4 & \tau_5 & \tau_4 & \tau_3 & \tau_1 & \tau_2 & \tau_1 & \tau_0 & \tau_4 & \tau_5 & \tau_4 & \tau_3 & \tau_7 & \tau_8 & \tau_7 & \tau_6 \\
\hline
\tau_6 & \tau_7 & \tau_8 & \tau_7 & \tau_3 & \tau_4 & \tau_5 & \tau_4 & \tau_0 & \tau_1 & \tau_2 & \tau_1 & \tau_3 & \tau_4 & \tau_5 & \tau_4 \\
\tau_7 & \tau_6 & \tau_7 & \tau_8 & \tau_4 & \tau_3 & \tau_4 & \tau_5 & \tau_1 & \tau_0 & \tau_1 & \tau_2 & \tau_4 & \tau_3 & \tau_4 & \tau_5 \\
\tau_8 & \tau_7 & \tau_6 & \tau_7 & \tau_5 & \tau_4 & \tau_3 & \tau_4 & \tau_2 & \tau_1 & \tau_0 & \tau_1 & \tau_5 & \tau_4 & \tau_3 & \tau_4 \\
\tau_7 & \tau_8 & \tau_7 & \tau_6 & \tau_4 & \tau_5 & \tau_4 & \tau_3 & \tau_1 & \tau_2 & \tau_1 & \tau_0 & \tau_4 & \tau_5 & \tau_4 & \tau_3 \\
\hline
\tau_3 & \tau_4 & \tau_5 & \tau_4 & \tau_6 & \tau_7 & \tau_8 & \tau_7 & \tau_3 & \tau_4 & \tau_5 & \tau_4 & \tau_0 & \tau_1 & \tau_2 & \tau_1 \\
\tau_4 & \tau_3 & \tau_4 & \tau_5 & \tau_7 & \tau_6 & \tau_7 & \tau_8 & \tau_4 & \tau_3 & \tau_4 & \tau_5 & \tau_1 & \tau_0 & \tau_1 & \tau_2 \\
\tau_5 & \tau_4 & \tau_3 & \tau_4 & \tau_8 & \tau_7 & \tau_6 & \tau_7 & \tau_5 & \tau_4 & \tau_3 & \tau_4 & \tau_2 & \tau_1 & \tau_0 & \tau_1 \\
\tau_4 & \tau_5 & \tau_4 & \tau_3 & \tau_7 & \tau_8 & \tau_7 & \tau_6 & \tau_4 & \tau_5 & \tau_4 & \tau_3 & \tau_1 & \tau_2 & \tau_1 & \tau_0
\end{array}\right).
\end{aligned}
$$

It turns out that the BCT-CT structure in (2.14) is a special case of the BCT structure where every matrix $\boldsymbol{\Sigma}_i$ in (2.12) is a $p \times p$ CT matrix with $[p/2]+1$ parameters, $i = 0, \ldots, [u/2]$. It has been shown by Nahtman and von Rosen (2008) that $\boldsymbol{\Sigma}_{BCT-CT}$ is invariant with respect to all orthogonal transformations $\boldsymbol{P}_1^{(1)} \otimes \boldsymbol{P}_2^{(1)}$, where $\boldsymbol{P}_1^{(1)}$ and $\boldsymbol{P}_2^{(1)}$ are two different SP matrices with sizes $u \times u$ and $p \times p$, respectively.

The study of the patterned covariance matrices with Kronecker structure $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$, where $\boldsymbol{\Sigma}(p \times p)$ and $\boldsymbol{\Psi}(q \times q)$, has raised much attention in recent years. Among others, this structure can be particularly useful to model spatial-temporal dependency simultaneously, where $\boldsymbol{\Sigma}$ is connected to temporal dependency and $\boldsymbol{\Psi}$ models the dependency over space (see Srivastava *et al.*, 2009, for example). From an inferential point of view, the Kronecker structure makes the estimation more complicated since the identification problem should be resolved and some restrictions have to be imposed on the parameter space. Then it results in non-explicit MLEs which depend on the choice of restrictions imposed on the covariance matrix (Srivastava *et al.*, 2008).

One interesting extension is when there can be some patterns imposed on matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$, e.g. the CS structure:

$$\begin{aligned}\boldsymbol{\Sigma}_{CS-CS} &= \boldsymbol{\Sigma}\otimes\boldsymbol{\Psi} = \big(a\boldsymbol{I}_p + b(\boldsymbol{J}_p - \boldsymbol{I}_p)\big)\otimes\big(c\boldsymbol{I}_q + d(\boldsymbol{J}_p - \boldsymbol{I}_q)\big),\\ &= \boldsymbol{I}_p\otimes a\big(c\boldsymbol{I}_q + d(\boldsymbol{J}_p - \boldsymbol{I}_q)\big) + (\boldsymbol{J}_p - \boldsymbol{I}_p)\otimes b\big(c\boldsymbol{I}_q + d(\boldsymbol{J}_p - \boldsymbol{I}_q)\big).\end{aligned}$$

Thus, it can be seen that $\boldsymbol{\Sigma}_{CS-CS}$ is also connected to the BCS-CS structure in (2.13).

3. Explicit maximum likelihood estimators in balanced models

One of the aims in this thesis is to discuss the existence of explicit MLEs of the (co)variance parameters for the random effects model presented in (1.2). Explicit estimators are often meaningful, because one can study basic properties of the estimators straightforwardly such as the distributions of estimators, without worrying about convergence problems as in the case of numerical estimation methods. In this chapter, the results derived by Szatrowski (1980) regarding the existence of explicit MLEs for both means and covariances in multivaraite normal models are presented. Szatrowski's results are applicable when the data is balanced, and in this thesis only balanced models are considered.

3.1 Explicit MLEs: Szatrowski's results

A result by Szatrowski, which provides the necessary and sufficient conditions for the existence of explicit MLEs for both means and (co)variance matrices with linear structures, can be applied in the context of the following general mixed linear model (Demidenko, 2004), of which model (1.2) is a special case:

$$\mathbf{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \tag{3.1}$$

where $\mathbf{y} : n \times 1$ is a response vector; matrices $\boldsymbol{X} : n \times m$ and $\boldsymbol{Z} : n \times q$ are known design and incidence matrices, respectively; $\boldsymbol{\beta} : m \times 1$ is a vector of fixed effects; $\boldsymbol{\gamma} : q \times 1$ is a vector of random effects; and $\boldsymbol{\epsilon} : n \times 1$ is a vector of random errors. Moreover, we assume that $E(\boldsymbol{\gamma}) = \mathbf{0}$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and

$$Var\begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{R} \end{pmatrix} = \begin{pmatrix} \boldsymbol{G} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R} \end{pmatrix},$$

where $\boldsymbol{G}$ is positive semidefinite and $\boldsymbol{R}$ is positive definite. Under a normality assumption on $\boldsymbol{\epsilon}$, we have $\mathbf{y} \sim N_n(\boldsymbol{X\beta}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = \boldsymbol{ZGZ}' + \boldsymbol{R}$ and $\boldsymbol{\Sigma}$ is assumed to be nonsingular. Usually, the term $\boldsymbol{Z\gamma}$ in (3.1) can be partitioned as

$$\boldsymbol{Z}\boldsymbol{\gamma} = (\boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_s) \begin{pmatrix} \boldsymbol{\gamma}_2 \\ \vdots \\ \boldsymbol{\gamma}_s \end{pmatrix}, \tag{3.2}$$

where $\boldsymbol{\gamma}_i$ can be a main effects factor, a nested factor or an interaction effects factor. Let n_i denote the number of levels of $\boldsymbol{\gamma}_i$. If the dispersion of $\boldsymbol{\gamma}_i$ is $\text{Var}(\boldsymbol{\gamma}_i) = \sigma_i^2 \boldsymbol{I}_{n_i}$, for all i, and $Cov(\boldsymbol{\gamma}_i, \boldsymbol{\gamma}'_h) = 0$, $i \neq h$, then

$$\boldsymbol{G} = \text{Diag}(\sigma_2^2 \boldsymbol{I}_{n_2}, \ldots, \sigma_s^2 \boldsymbol{I}_{n_s}),$$

and $\boldsymbol{R} = \sigma_1^2 \boldsymbol{I}_n$ may also be assumed. Define $\boldsymbol{\gamma}_1 = \boldsymbol{\epsilon}$, $n_1 = n$ and $\boldsymbol{Z}_1 = \boldsymbol{I}_n$. The covariance matrix of $\mathbf{y}$ can be written as a linear structure in (2.1), i.e. $\boldsymbol{\Sigma} = \sum_{i=1}^{s} \theta_i \boldsymbol{V}_i$, where $\boldsymbol{V}_i = \boldsymbol{Z}_i \boldsymbol{Z}'_i$. Since $\boldsymbol{\Sigma}$ is a function of $\boldsymbol{\theta}$, it is denoted by $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ comprise all unknown parameters in the matrices $\boldsymbol{G}$ and $\boldsymbol{R}$.

In practice, the estimation of both $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ is of primary interest. Several estimation methods can be used, e.g. ML estimation and REML estimation which both rely on the normal distributional assumption, analysis of variance estimation (ANOVA) and minimum norm quadratic unbiased estimation (MINQUE). We may also use Bayesian estimation, which starts with prior distributions for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ and results in a posterior distribution of the unknown parameters after observing the data.

The likelihood function for $\mathbf{y}$, which is the function of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ equals

$$L(\boldsymbol{\beta}, \boldsymbol{\theta} | \mathbf{y}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}(\boldsymbol{\theta})|^{-1/2} exp\left[-(\mathbf{y} - \boldsymbol{X}\boldsymbol{\beta})' \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} (\mathbf{y} - \boldsymbol{X}\boldsymbol{\beta})/2 \right],$$

where $|\bullet|$ denotes the determinant of a matrix. Let $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ denote the MLE of $\boldsymbol{X}\boldsymbol{\beta}$. Using the normal equation $\boldsymbol{X}'\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\mathbf{y}$, we have

$$\boldsymbol{X}\hat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})^{-1}\mathbf{y}, \tag{3.3}$$

where $\hat{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$.

For (3.3), several authors have discussed the conditions of loosening dependence on $\boldsymbol{\theta}$ (and hence $\hat{\boldsymbol{\theta}}$) in $\boldsymbol{X}\hat{\boldsymbol{\beta}}$; for example, see Zyskind (1967), Mitra and Moore (1973) and Puntanen and Styan (1989). If $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ does not depend on $\boldsymbol{\theta}$, then $\boldsymbol{X}\hat{\boldsymbol{\beta}}$ results in an ordinary least square estimator (OLS) in model (3.1).

According to the result in Szatrowski (1980), a necessary and sufficient condition for

$$(\boldsymbol{X}'\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})^{-1} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$$

is that there exists a subset of r orthogonal eigenvectors of $\boldsymbol{\Sigma}$ which form a basis of $\mathcal{C}(\boldsymbol{X})$, where $r = \text{rank}(\boldsymbol{X})$ and $\mathcal{C}(\bullet)$ denotes the column vector space.

Alternatively, one can state that $\mathcal{C}(\boldsymbol{X})$ has to be $\boldsymbol{\Sigma}$-invariant in order to obtain explicit estimators, i.e. $\boldsymbol{\beta}$ in (3.1) has explicit MLE if and only if $\mathcal{C}(\boldsymbol{\Sigma X}) \subseteq \mathcal{C}(\boldsymbol{X})$. Shi and Wang (2006) obtained an equivalent condition, namely $\boldsymbol{P_X \Sigma}$ should be symmetric, where $\boldsymbol{P_X} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}$.

In the context of the growth curve model (Kollo and von Rosen, 2005, Chapter 4), Rao (1967) have showed that for certain covariance structures, the unweighted estimator (LSE) for the mean is the MLE. This fact was presented by Puntanen and Styan (1989) as an example. Consider the following mixed model:

$$\mathbf{y} = \boldsymbol{X\beta} + \boldsymbol{X\gamma} + \boldsymbol{Z\xi} + \boldsymbol{\epsilon}, \tag{3.4}$$

where $\boldsymbol{Z}$ is a matrix such that $\boldsymbol{X}'\boldsymbol{Z} = 0$, $\boldsymbol{\gamma}$, $\boldsymbol{\xi}$ and $\boldsymbol{\epsilon}$ are uncorrelated random vectors with zero expectations and covariance matrices $\boldsymbol{\Gamma}$, $\boldsymbol{C}$ and $\sigma^2\boldsymbol{I}$, respectively. In model (3.4) the covariance matrix of $\mathbf{y}$ belongs to the class of so-called Rao's simple covariance structure (Pan and Fang, 2002), i.e.,

$$\mathrm{Var}(\mathbf{y}) = \boldsymbol{X\Gamma X}' + \boldsymbol{ZCZ}' + \sigma^2\boldsymbol{I}.$$

Now we are going to present Szatrowski's result of explicit MLEs for (co)variance parameters. The result assumes that the covariance matrix satisfies a canonical form, i.e. there exists a value $\boldsymbol{\theta}^* \in \boldsymbol{\Theta}$ such that $\boldsymbol{\Sigma}(\boldsymbol{\theta}^*) = \boldsymbol{I}$, where $\boldsymbol{\Theta}$ represents the parameter space, or can be transformed into this form. Moreover, the following result given by Roebruck (1982) indicates that the study of the spectral decomposition (or eigen-decomposition) of patterned covariance matrices is crucial when finding explicit MLEs of the covariances.

Theorem 3.1.1 *(Roebruck, 1982, Theorem 1) Assume that the matrix $\boldsymbol{X}$ is of full column rank m. Model* (3.1) *has a canonical form if and only if there exists a set of n linearly independent eigenvectors of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, which are independent of $\boldsymbol{\theta}$ and m of which span the column space of $\boldsymbol{X}$.*

The following theorem provides necessary and sufficient conditions for the existence of explicit MLEs for the (co)variance parameters $\boldsymbol{\theta}$.

Theorem 3.1.2 *(Szatrowski, 1980) Assume that the MLE of $\boldsymbol{\beta}$ has an explicit representation and that $\boldsymbol{V}$'s in $\boldsymbol{\Sigma} = \sum_{i=1}^{s} \theta_i \boldsymbol{V}_i$ are all diagonal in the canonical form. Then, the MLE of $\boldsymbol{\theta}$ has an explicit representation if and only if the diagonal elements of $\boldsymbol{\Sigma}$ consist of exactly s linearly independent combinations of $\boldsymbol{\theta}$.*

Note that $\boldsymbol{\Sigma}$ in Theorem 3.1.2 is diagonal due to the spectral decomposition. Hence, the diagonal elements of $\boldsymbol{\Sigma}$ are actually the eigenvalues of the original covariance matrix. Theorem 3.1.2 is essential when studying explicit MLEs of (co)variance parameters and hence has been referred to several times in this thesis (Papers II-III). Illustrations of this result as well as discussions can be found in (Szatrowski and Miller, 1980). For inference in unbalanced mixed models, for example, see Jennrich and Schluchter (1986), which described Newton-Raphson and Fisher scoring algorithms for computing MLEs of $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$, and generalized EM algorithms for computing restricted and unrestricted MLEs.

3.2 Spectral decomposition of pattern covariance matrices

The importance of the spectral decomposition when making inference for patterned covariance matrices has been noticed in many previous studies (see Olkin and Press, 1969; Arnold, 1973; Krishnaiah and Lee, 1974; Szatrowski and Miller, 1980, for example). In this section we summarize the spectral decompositions for different block covariance structures that are used to derive explicit estimators. To be more accurate, here the term "spectral decomposition" means not only eigenvalue decomposition but also eigenblock (eigenmatrix) decomposition. The following eigenvalues or eigenblocks can be considered as the reparametrization of the original block structures and they are one-to-one transformations of the parameter spaces, which play an important role in both estimation and construction of likelihood ratio tests (see Chapter 4).

In order to present the results we will first define two orthogonal matrices that will be used in the following various spectral decompositions. Let $\boldsymbol{K}$ be a Helmert matrix, i.e. an $u \times u$ orthogonal matrix such that

$$\boldsymbol{K}_u = (u^{-1/2}\mathbf{1}_u \vdots \boldsymbol{K}_1), \tag{3.5}$$

where $\boldsymbol{K}_1'\mathbf{1}_u = \mathbf{0}$ and $\boldsymbol{K}_1'\boldsymbol{K}_1 = \boldsymbol{I}_{u-1}$. Let $\boldsymbol{V}$ be another $p \times p$ orthogonal matrix such that

$$\boldsymbol{V}_p = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^p), \tag{3.6}$$

where the vectors $\boldsymbol{v}^1, \ldots, \boldsymbol{v}^p$ are the orthonormal eigenvectors of the CT matrix in (2.3). For the derivation of the matrix $\boldsymbol{V}_p$, we refer readers to Basilevsky (1983).

The CS matrix of size $p \times p$ in (2.2) can be decomposed as

$$\boldsymbol{\Sigma}_{CS} = \boldsymbol{K}_p \mathrm{Diag}(\boldsymbol{\lambda}) \boldsymbol{K}_p',$$

where $\text{Diag}(\boldsymbol{\lambda})$ is a diagonal matrix with the diagonal elements $a+(p-1)b$ or $a-b$, i.e. the eigenvalues of the CS matrix. The CT matrix in (2.3) can be decomposed as $\boldsymbol{\Sigma}_{CT} = \boldsymbol{V}_p \text{Diag}(\boldsymbol{\lambda}) \boldsymbol{V}'_p$, where $\text{Diag}(\boldsymbol{\lambda})$ is a diagonal matrix with the diagonal elements

$$\lambda_k = \sum_{j=0}^{p-1} t_j \cos\left(\frac{2\pi}{p}(k-1)(p-j)\right), \quad k = 1, \ldots, p, \tag{3.7}$$

where t_j is the element of $\boldsymbol{\Sigma}_{CT}$ in (2.3).

In Chapter 2, we presented different block covariance structures as well as their potential utilization. Now the spectral decompositions of those structures will be given, and the results are crucial from an inferential point of view. The matrix in (2.9) can be block-diagonalized as follows (Arnold, 1979):

$$(\boldsymbol{K}'_u \otimes \boldsymbol{I}_p)\boldsymbol{\Sigma}_{BCS}(\boldsymbol{K}_u \otimes \boldsymbol{I}_p) = \begin{pmatrix} \boldsymbol{\Sigma}_0 + (u-1)\boldsymbol{\Sigma}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{u-1} \otimes (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1) \end{pmatrix}, \tag{3.8}$$

where $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ are the matrices given in (2.9). Here the matrices $\boldsymbol{\Sigma}_0 + (u-1)\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1$ are called eigenblocks.

The matrix in (2.13) can be diagonalized as follows (Nahtman, 2006):

$$(\boldsymbol{K}'_u \otimes \boldsymbol{V}'_p)\boldsymbol{\Sigma}_{BCS-CS}(\boldsymbol{K}_u \otimes \boldsymbol{V}_p) = \text{Diag}(\boldsymbol{\lambda}), \tag{3.9}$$

where $\boldsymbol{K}_u$ is given in (3.5), $\boldsymbol{V}_p$ is given in (3.6), and $\text{Diag}(\boldsymbol{\lambda})$ is a $up \times up$ diagonal matrix with elements

$$\begin{aligned} \lambda_1 &= a+(p-1)b+(u-1)\left[c+(p-1)d\right], \\ \lambda_2 &= a-b+(u-1)(c-d), \\ \lambda_3 &= a+(p-1)b-\left[c+(p-1)d\right], \\ \lambda_4 &= a-b-(c-d), \end{aligned}$$

of multiplicity

$$m_1 = 1, \quad m_2 = p-1, \quad m_3 = u-1 \quad \text{and} \quad m_4 = (u-1)(p-1),$$

respectively. It is seen from (3.9) that the eigenvalues of $\boldsymbol{\Sigma}_{BCS-CS}$ can be expressed as linear combinations of the eigenvalues of the blocks, when $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ are CS structures.

The matrix in (2.10) can be diagonalized as follows:

$$(\boldsymbol{K}'_u \otimes \boldsymbol{V}'_p)\boldsymbol{\Sigma}_{DCS}(\boldsymbol{K}_u \otimes \boldsymbol{V}_p) = \text{Diag}(\boldsymbol{\lambda}),$$

where $\boldsymbol{K}_u$ is given in (3.5), $\boldsymbol{V}_p$ is given in (3.6), and $\text{Diag}(\boldsymbol{\lambda})$ is a $up \times up$ diagonal matrix with the elements

$$\begin{aligned} \lambda_1 &= a - b + p(b-c) + puc, \\ \lambda_2 &= a - b, \\ \lambda_3 &= a - b + p(b-c), \end{aligned} \tag{3.10}$$

of multiplicity

$$m_1 = 1, \quad m_2 = u(p-1) \quad \text{and} \quad m_3 = u-1,$$

respectively. Additionally, we have the restriction $c < b - \frac{b-a}{p}$ to preserve the positive definiteness of $\boldsymbol{\Sigma}_{DCS}$.

The block diagonalization of the matrix $\boldsymbol{\Sigma}_{BDCS}$ in (2.11) refers to the result of Roy and Fonseca (2012), and it has the following three distinct eigenblocks:

$$\begin{aligned} \boldsymbol{\Lambda}_1 &= (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1) + u(\boldsymbol{\Sigma}_1 - \boldsymbol{W}) + uv\boldsymbol{W}, \\ \boldsymbol{\Lambda}_2 &= \boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1, \\ \boldsymbol{\Lambda}_3 &= (\boldsymbol{\Sigma}_0 - \boldsymbol{\Sigma}_1) + u(\boldsymbol{\Sigma}_1 - \boldsymbol{W}), \end{aligned} \tag{3.11}$$

of multiplicity 1, $v(u-1)$ and $v-1$, respectively. Comparing (3.11) and (3.10), similar structures can be observed, and (3.11) will degenerate to (3.10) when both $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Sigma}_1$ are two different scalars instead of matrices.

The matrix in (2.12) can be block-diagonalized as follows (Olkin, 1973b):

$$(\boldsymbol{V}'_u \otimes \boldsymbol{I}_p)\boldsymbol{\Sigma}_{BCT}(\boldsymbol{V}_u \otimes \boldsymbol{I}_p) = \text{Diag}(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_u), \tag{3.12}$$

where $\text{Diag}(\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_u)$ is a block diagonal matrix with the matrices $\boldsymbol{\psi}_j$ which are positive definite and satisfy $\boldsymbol{\psi}_j = \boldsymbol{\psi}_{u-j+2}$, $j = 2, \ldots, u$.

The matrix in (2.14) can be diagonalized as follows (Nahtman and von Rosen, 2008):

$$(\boldsymbol{V}'_u \otimes \boldsymbol{V}'_p)\boldsymbol{\Sigma}_{BCT-CT}(\boldsymbol{V}_u \otimes \boldsymbol{V}_p) = \sum_{k_2=0}^{[u/2]} \text{Diag}_{k_2}(\boldsymbol{\lambda}) \otimes \text{Diag}_{CT,k_2}(\boldsymbol{\lambda}), \tag{3.13}$$

where $\text{Diag}_{k_2}(\boldsymbol{\lambda})$ is a diagonal matrix with the diagonal elements are the eigenvalues of the symmetric circular matrix $SC(u, k_1)$ (as a special case of the CT matrix) in (2.14), and $\text{Diag}_{CT,k_2}(\boldsymbol{\lambda})$ is another diagonal matrix with the diagonal elements the eigenvalues of the CT matrix $\sum_{k_1=0}^{[p/2]} t_k SC(p, k_1)$, where $k = (\frac{p}{2} + 1)k_2 + k_1$.

Here a similar relationship between (2.9) and (2.13) can be observed when comparing with (3.12) and (3.13). The eigenvalues of $\boldsymbol{\Sigma}_{BCT-CT}$ is expressed as linear combinations of the eigenvalues of the blocks $\boldsymbol{\psi}_j$ when $\boldsymbol{\psi}_j$ in (3.12) has the CT structures, $j = 1, \ldots, u$.

Seen from the spectral decompositions above, the patterned matrices are either diagonalized or block-diagonalized by the orthogonal matrices, which are not a function of the elements in those matrices, and which will be very useful when connecting with other covariance structures, deriving likelihood ratio tests as well as studying their corresponding distributions. In this thesis, the spectra of our new block covariance structures have also been obtained in a similar way, see the summary of Papers I-II in Chapter 5.

4. Testing block covariance structures

It is very often necessary to check whether the assumptions imposed on various covariance matrices are satisfied. Testing the validity of covariance structures is crucial before using them for any statistical analysis. Paper IV in this thesis focuses on developing LRT procedures for testing certain block covariance structures, as well as the (co)variance parameters of the block circular Toeplitz structure. In this chapter we focus on the introduction of the likelihood ratio test (LRT) procedure together with the approximations of the null distributions of the LRT statistic following Box (1949).

4.1 Likelihood ratio test procedures for testing covariance structures

4.1.1 Likelihood ratio test

LRT plays an important role in testing certain hypotheses on mean vectors and covariance matrices under various model settings, for example ANOVA and MANOVA models (Krishnaiah and Lee, 1980). This regards an LRT criterion Λ for testing the mean $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ under the *null hypothesis* $H_0 : \Theta_0$ versus the *alternative hypothesis* $H_a : \Theta$, assuming the restricted parameter space $\Theta_0 \subset \Theta$, is constructed by

$$\Lambda = \frac{\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}\in\Theta_0} L(\boldsymbol{\mu},\boldsymbol{\Sigma})}{\max_{\boldsymbol{\mu},\boldsymbol{\Sigma}\in\Theta} L(\boldsymbol{\mu},\boldsymbol{\Sigma})},$$

where max is the maximization function. The null hypothesis H_0 is rejected if $\Lambda \leq c$, where c is chosen such that the significance level is α. It is well known that under the null hypothesis H_0, the quantity $-2\ln\Lambda$ is asymptotically χ^2 distributed with degrees of freedom equal to the difference in the dimensionality of Θ_0 and Θ.

When the multivariate normality assumption is assumed, there is a comprehensive study of likelihood ratio procedures for testing the hypotheses of the equality of covariance matrices, and the equality of both covariance matrices and mean vectors (e.g. see Anderson, 2003, Chapter 10). The study of testing the block CS covariance matrix can be traced back to Votaw (1948). He extended the testing problem of CS structure (Wilks, 1946) to the "block version" and developed LRT criteria for testing 12 hypotheses, e.g. the hypothesis of the equality of means, the equality of variances and the equality of covariances, which were applied to certain psychometric and medical research problems. Later Olkin (1973b) considered the problem of testing the circular Toeplitz covariance matrix in blocks, which is also an "block" extension of the previous work by Olkin and Press (1969).

Besides LRT, Rao's score test (RST) has also been discussed in the literature, and for RST we only need to exploit the null hypothesis, i.e. calculate the score vector and Fisher information matrix evaluated at the MLEs under the null hypothesis. Chi and Reinsel (1989) derive RST for a AR(1) structure. Computationally intensive procedures for testing covariance structures have also been developed, such as parametric bootstrap tests and permutation tests.

4.1.2 Null distributions of the likelihood ratio test statistics and Box's approximation

As mentioned above, it is well known that the asymptotic null distribution of $-2\ln\Lambda$ is a χ^2-distribution with degrees of freedom equal to the difference in dimensionality of Θ and Θ_0, see Wilks (1938), for example.

However, in many situations with small sample sizes, the asymptotic χ^2 distribution is not an adequate approximation. One way to improve the χ^2 approximation of the LRT statistic is the Box's approximation. Box (1949) provided an approximate null distribution of $-2\ln\Lambda$ in terms of a linear combination of central χ^2 distributions. Once the moments of the LRT statistic Λ $(0 \leq \Lambda \leq 1)$ is derived in terms of certain functions of Gamma functions, then Box's approximation can be applied. The result of Box can be expressed as follows:

Theorem 4.1.1 *(Anderson, 2003, p.316) Consider a random variable* $\Lambda(0 \leq \Lambda \leq 1)$ *with* s*-th moment*

$$E(\Lambda^s) = K\left[\frac{\prod_{j=1}^{b} y_j^{y_j}}{\prod_{k=1}^{a} x_k^{x_k}}\right]^s \frac{\prod_{k=1}^{a}\Gamma[x_k(1+s)+\delta_k]}{\prod_{j=1}^{b}\Gamma[y_j(1+s)+\eta_j]}, \quad s=0,1,\ldots,$$

where K *is a constant such that* $E(\Lambda^0)=1$ *and* $\sum_{k=1}^{a} x_k = \sum_{j=1}^{b} y_j$. *Then,*

$$P(-2\rho\ln\Lambda \leq t) = P(\chi_f^2 \leq t) + O(n^{-2}),$$

where $O(n^{-2})$ denotes any quantity that if there exist M and n_0, $|O(n^{-2})/n^{-2}| < M$ for all $n > n_0$,

$$f = -2\left[\sum_{k=1}^{a}\delta_k - \sum_{j=1}^{b}\eta_j - \frac{a-b}{2}\right],$$

and ρ is the solution of

$$\sum_{k=1}^{a}\frac{B_2(\beta_k+\delta_k)}{x_k} = \sum_{j=1}^{b}\frac{B_2(\epsilon_j+\eta_j)}{y_j},$$

where $\beta_k = (1-\rho)x_k$, $\epsilon_j = (1-\rho)y_j$ and B_2 is the Bernoulli polynomial of degree 2, i.e. $B_2(x) = x^2 - x + 1/6$.

Many LRT statistics concerning testing in multivariate normal models have the moments expressed in the form given of Theorem 4.1.1 and they have the null distributions expressed in terms of the products of independent beta random variables, for example, when testing the equality of several mean vectors, the equality of several covariance matrices or the sphericity of the covariance matrix (Muirhead, 1982; Anderson, 2003) as well as testing circularity of the covariance matrix (Olkin and Press, 1969; Olkin, 1973b).

4.2 F test and likelihood ratio test of variance components

Exact tests for testing variance components started from Wald (1941, 1947). He derived exact tests for one-way and two-way cross-classification models without interactions. Seely and El-Bassiouni (1983) considered extensions of Wald's variance component test in context of ordinary mixed linear models and provided necessary and sufficient conditions for the test proposed by Wald to be applicable. Later Gallo and Khuri (1990) presented exact tests concerning the variance components in the unbalanced two-way cross-classification model. Öfversten (1993) presented two kinds of exact F-tests for variance components in unbalanced mixed linear models for which derivation was based on a preliminary orthogonal transformation and a subsequent resampling procedure.

It has been noticed that zero-variance hypothesis is not a standard testing problem since the hypothesis is on the boundary of the parameter space. Self and Liang (1987) derived a large sample mixture of chi-square distributions of LRT using the usual asymptotic theory for a null hypothesis on the boundary of the parameter space. Crainiceanu and Ruppert (2004) developed finite samples and asymptotic distributions for both LRT and restricted LRT concerning mixed linear models with one variance component.

Some of the hypotheses of testing (co)variance parameters in Paper IV of this thesis are also on the boundary of the parameter space. The tests we have constructed which are based on the likelihood ratios, however, do not need any restrictions on the parameter space of the (co)variance parameters. Srivastava and Singull (2012) considered hypothesis testing for a parallel profile model with a CS random-effects covariance structure $\mathbf{\Sigma}_{CS}$, given in (2.2), and it has been clarified that only the distinct eigenvalues of $\mathbf{\Sigma}_{CS}$ are necessary to be estimated rather than the original (co)variance parameters. Moreover, Srivastava and Singull (2012) concluded that the restriction of the positiveness of the variance parameter is unnecessary when dealing with hypothesis testing.

For some of the testing problems for the (co)variance parameters considered in this thesis, it can be shown that to test each hypothesis of interest requires nothing but testing the equality of several variances. In this case, we can rely on existing methods such as Bartlett's test (Bartlett, 1937). However, there are some tests where the testability (identifiability) problem has to be investigated carefully before a test can be constructed, see the summary of Paper IV in Chapter 5.

5. Summary of papers

The results in this thesis consist of the derivation of specific block covariance structures and more importantly, the inferential results of multivariate normal models with block circular Toeplitz structures. In this chapter, the main results will be highlighted across relevant sections.

5.1 Paper I: Block circular symmetry in multilevel models

Compound symmetry and circular symmetry are two different ways to model data. Considering the situations of Examples 4 and 5, when they appear simultaneously, what is the corresponding covariance structure in order to characterize this type of dependency?

Paper I deals with a particular class of covariance matrices that are invariant under two types of orthogonal transformations, $\boldsymbol{P}^{(2)} \otimes \boldsymbol{P}^{(1)}$ and $\boldsymbol{P}^{(1)} \otimes \boldsymbol{P}^{(2)}$, where $\boldsymbol{P}^{(2)}$ is any permutation matrix and $\boldsymbol{P}^{(1)}$ is any shift-permutation matrix given in (2.7). It was shown that the two orthogonal actions imply two different block symmetric covariance structures. The following necessary and sufficient conditions reveal the corresponding covariance structures.

Theorem 5.1.1 *(Theorem 3.3, Paper I, p.10) The covariance matrix $\boldsymbol{\Sigma}_{21}$: $n_2 n_1 \times n_2 n_1$ is invariant with respect to all orthogonal transformations defined by $\boldsymbol{P}_{21} = \boldsymbol{P}^{(2)} \otimes \boldsymbol{P}^{(1)}$, if and only if it has the following structure:*

$$\boldsymbol{\Sigma}_{21} = \boldsymbol{I}_{n_2} \otimes \sum_{k_1=0}^{[n_1/2]} \tau_{k_1} SC(n_1, k_1) + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \sum_{k_1=0}^{[n_1/2]} \tau_{k_1+[n_1/2]+1} SC(n_1, k_1), \quad (5.1)$$

where τ_{k_1} and $\tau_{k_1+[n_1/2]+1}$ are constants, and the matrices $SC(n_1, k_1)$ are symmetric circular matrices defined in (2.5)*, $k_1 = 0, \dots, [n_1/2]$.*

Theorem 5.1.2 *(Theorem 3.5, Paper I, p.14) The covariance matrix $\boldsymbol{\Sigma}_{12}$: $n_2 n_1 \times n_2 n_1$ is invariant with respect to all orthogonal transformations defined by $\boldsymbol{P}_{12} = \boldsymbol{P}^{(1)} \otimes \boldsymbol{P}^{(2)}$ if and only if it has the following structure:*

$$\boldsymbol{\Sigma}_{12} = \sum_{k_2=0}^{[n_2/2]} \left[SC(n_2, k_2) \otimes \boldsymbol{\Sigma}^{(k_2)} \right], \quad (5.2)$$

where $\boldsymbol{\Sigma}^{(k_2)} = \tau_{k_2}\boldsymbol{I}_{n_1} + \tau_{k_2+[n_2/2]+1}(\boldsymbol{J}_{n_1} - \boldsymbol{I}_{n_1})$, τ_{k_2} *and* $\tau_{k_2+[n_2/2]+1}$ *are constants.* $SC(n_2, k_2)$ *is the symmetric circular matrix given in* (2.5).

The results given above are useful for characterization of the dependency in the context of multivariate two-level data. The structure $\boldsymbol{\Sigma}_{21}$ in (5.1) extends the covariance structures given by (2.9) and considers CT structures in each block, while the structure $\boldsymbol{\Sigma}_{12}$ in (5.2) is an extension of the structure in (2.12) when the CS structure is imposed in each block. The structures are also called mixed block structures, and the terminology “mixed block" refers to combining two different invariant properties, which was introduced by Barton and Fuhrmann (1993) when describing the dependency of array signal processing data.

Moreover, in Paper I we also demonstrated the relationship between these two structures by utilizing the commutation matrix (Kollo and von Rosen, 2005, Definition 1.3.2, p.79), which is used to relabel observations, see Theorem 3.7, Paper I, p.17. This simplifies the situation when discussing estimation of the model parameters since it is enough to consider only one covariance structure, which will be $\boldsymbol{\Sigma}_{21}$ in the follow-up statistical inferential studies, Papers II-IV. It is worth noting that even though this thesis is interested in multivaraite two-level data, it does not have to be nested in order to use the derived covariance structures.

The spectra, i.e. the set of eigenvalues of the two types of block circular symmetric covariance matrices are also obtained. Moreover, it can be seen that the two matrices $\boldsymbol{\Sigma}_{21}$ and $\boldsymbol{\Sigma}_{12}$ have the same spectra since they are similar matrices. The spectral property of the covariance matrix given in Theorem 5.1.1 can be derived directly by using the following theorem.

Theorem 5.1.3 *(Theorem 4.1, Paper I, p.21) Let the covariance matrix* $\boldsymbol{\Sigma}_{21} : n_2n_1 \times n_2n_1$ *have the structure obtained in Theorem 5.1.1. Let* $\lambda_h^{(i)}$ *be the eigenvalue of* $\boldsymbol{\Sigma}^{(i)} : n_1 \times n_1$ *with multiplicity* m_h, $i, = 1, 2, h = 1, \ldots, [n_1/2] + 1$. *The spectrum of* $\boldsymbol{\Sigma}_{21}$ *consists of the eigenvalues* $\lambda_h^{(1)} - \lambda_h^{(2)}$, *each of multiplicity* $(n_2 - 1)m_h$, *and* $\lambda_h^{(1)} + (n_2 - 1)\lambda_h^{(2)}$, *each of multiplicity* m_h. *The number of distinct eigenvalues is* $2([n_1/2] + 1)$.

The novelty of our results concerning the spectra of block circular symmetric matrices is that the eigenvalues of these block matrices can be expressed as linear combinations of the eigenvalues of the blocks instead of direct calculations using the matrix elements. The provided results describe the eigenvalues of patterned covariance matrices in a systematic way. During the proof, we use such properties as *commutativity* and *simultaneous diagonalization*, i.e. if two normal matrices commute then they have a joint eigenspace and can be diagonalized simultaneously. The multiplicities of eigenvalues, and the number of distinct eigenvalues of the two types of patterned covariance structures presented in Theorems 5.1.1-5.1.2, are also given.

5.2 Paper II: On estimation in multilevel models with block circular symmetric covariance structure

Paper II considers the MLE of parameters in model (1.2) when the covariance matrix $\boldsymbol{\Sigma}$ in (1.3) is block circular symmetric with CS patterned blocks. As noted in Chapter 1, this covariance structure can be used to characterize data with the features of circularity and exchangeability. The derived results can be considered as a complement to earlier works (Olkin and Press, 1969; Olkin, 1973b) in the sense of studying the estimation of a new type of multivariate two-level data together with circular symmetric models.

Recall the covariance matrix $\boldsymbol{\Sigma}$ in model (1.2). The next example illustrates a block covariance structure $\boldsymbol{\Sigma}$ in (1.3) when $n_2 = 3$ and $n_1 = 4$, where n_2 and n_1 are the number of factor levels for $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$, respectively.

Example 6 *Suppose a measurement is made at each of n_1 factor levels among n_2 factor levels and assume there are n independent units available. When $n_2 = 3$ and $n_1 = 4$, for each unit, the covariance matrix $\boldsymbol{\Sigma} : 12 \times 12$ in (1.3) has the following form:*

$$\boldsymbol{\Sigma} = \boldsymbol{I}_3 \otimes \begin{pmatrix} \sigma^2+\sigma_1+\tau_1 & \sigma_1+\tau_2 & \sigma_1+\tau_3 & \sigma_1+\tau_2 \\ \sigma_1+\tau_2 & \sigma^2+\sigma_1+\tau_1 & \sigma_1+\tau_2 & \sigma_1+\tau_3 \\ \sigma_1+\tau_3 & \sigma_1+\tau_2 & \sigma^2+\sigma_1+\tau_1 & \sigma_1+\tau_2 \\ \sigma_1+\tau_2 & \sigma_1+\tau_3 & \sigma_1+\tau_2 & \sigma^2+\sigma_1+\tau_1 \end{pmatrix}$$
$$+ (\boldsymbol{J}_3 - \boldsymbol{I}_3) \otimes \begin{pmatrix} \sigma_2+\tau_4 & \sigma_2+\tau_5 & \sigma_2+\tau_6 & \sigma_2+\tau_5 \\ \sigma_2+\tau_5 & \sigma_2+\tau_4 & \sigma_2+\tau_5 & \sigma_2+\tau_6 \\ \sigma_2+\tau_6 & \sigma_2+\tau_5 & \sigma_2+\tau_4 & \sigma_2+\tau_5 \\ \sigma_2+\tau_5 & \sigma_2+\tau_6 & \sigma_2+\tau_5 & \sigma_2+\tau_4 \end{pmatrix}, \qquad (5.3)$$

where the diagonal blocks represent the 4×4 variances and covariances of the 4 measurements coming from the same level of $\boldsymbol{\gamma}_1$ and the off-diagonal blocks represent the 4×4 covariances of the 4 measurements between any pair of levels of $\boldsymbol{\gamma}_1$.

The spectral properties of block circular symmetric covariance matrix $\boldsymbol{\Sigma}$ with patterned blocks are derived. We also give the actual number of all distinct eigenvalues and their expressions. The covariance matrix $\boldsymbol{\Sigma}$ of model (1.2) given in (1.3) is a sum of three symmetric matrices $\boldsymbol{Z}_1\boldsymbol{\Sigma}_1\boldsymbol{Z}_1'$, $\boldsymbol{\Sigma}_2$ and $\sigma^2\boldsymbol{I}_p$, which has been shown to commute (see Lemma 2.1, Paper II), and hence can be simultaneously diagonalized. This fact is utilized to obtain the eigenvalues of $\boldsymbol{\Sigma}$ which are presented in the next theorem.

Theorem 5.2.1 *(Theorem 3.1, Paper II, p.89) Let the matrix $\boldsymbol{\Sigma}$ be defined as in* (1.3). *There exists an orthogonal matrix $\boldsymbol{Q} = \boldsymbol{K} \otimes \boldsymbol{V}$ such that $\boldsymbol{Q}'\boldsymbol{\Sigma}\boldsymbol{Q} = \boldsymbol{D}$, where $\boldsymbol{K}$ and $\boldsymbol{V}$ are defined in* (3.5) *and* (3.6)*, respectively, and $\boldsymbol{D}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\Sigma}$. Moreover,*

$$\boldsymbol{D} = Diag\big(\boldsymbol{D}_1, \boldsymbol{I}_{n_2-1} \otimes \boldsymbol{D}_2\big),$$

where

$$\begin{aligned}
\boldsymbol{D}_1 &= Diag(\sigma^2 + n_1 a + n_1(n_2-1)b + \lambda_{11}, \sigma^2 + \lambda_{12}, \ldots, \sigma^2 + \lambda_{1n_1}),\\
\boldsymbol{D}_2 &= Diag(\sigma^2 + n_1(a-b) + \lambda_{21}, \sigma^2 + \lambda_{22}, \ldots, \sigma^2 + \lambda_{2n_1}),
\end{aligned}$$

and λ_{ih} are the eigenvalues given in Theorem 5.1.3, $i = 1,2$, $h = 1,\ldots,n_1$.

By using the spectral decomposition, it is shown to be a covariance matrix with a linear structure (Anderson, 1973). The spectral decomposition of $\boldsymbol{\Sigma}$ has been utilized to obtain explicit MLEs for the mean parameter μ and the covariance matrix $\boldsymbol{\Sigma}$. Recall the mean structure of model (1.2), i.e. $\mathbf{1}_p\mu$, and we have that $\mathcal{C}(\boldsymbol{\Sigma}\mathbf{1}_p) = \mathcal{C}(\mathbf{1}_p)$ holds. According to the result presented in Szatrowski (1980), the MLE of μ is just the average of the total np observations. The MLE for $\boldsymbol{\Sigma}$ has been derived through the MLEs of the distinct eigenvalues of $\boldsymbol{\Sigma}$, see Theorem 4.1, Paper II, p.93.

Under the existence of the explicit MLE of μ, our main concern is the existence of the explicit MLE of the (co)variance parameters contained in $\boldsymbol{\Sigma}$, denoted as $\boldsymbol{\theta}$. According to Theorem 3.1.2, it is noted that explicit MLEs for r (co)variance parameters in the balanced linear model exist if and only if all distinct eigenvalues of $\boldsymbol{\Sigma}$ are r linearly independent combinations of (co)variance parameters. We proved that the difference between the number of distinct eigenvalues of $\boldsymbol{\Sigma}$ and the number of unknown (co)variance parameters equals 3, i.e.,

$$\boldsymbol{\Sigma} = \underbrace{\boldsymbol{Z}_1\boldsymbol{\Sigma}_1\boldsymbol{Z}_1'}_{\text{2 parameters}} + \underbrace{\boldsymbol{\Sigma}_2}_{2r\text{ parameters}} + \underbrace{\sigma^2\boldsymbol{I}}_{\text{1 parameter}}. \tag{5.4}$$

Thus, there are $2r+3$ unknown parameters in $\boldsymbol{\Sigma}$, whereas there are only $2r$ distinct eigenvalues of $\boldsymbol{\Sigma}$ (see Table 5.1), where $r = [n_1/2]+1$ and $[\bullet]$ denotes the integer function. Therefore, explicit MLEs for all (co)variance parameters do not exist in the considered model.

Table 5.1. *Distinct eigenvalues η_i of $\boldsymbol{\Sigma}$ given in (1.3) with corresponding multiplicities m_i.*

	m_i	
η_i	odd n_1	even n_1
η_1	1	1
$\eta_2,\ldots,\eta_{[\frac{n_1}{2}]+1}$	2	2, $\eta_{\frac{n_1}{2}}$ has multiplicity 1.
$\eta_{[\frac{n_1}{2}]+2}$	n_2-1	n_2-1
$\eta_{[\frac{n_1}{2}]+3},\ldots,\eta_{2([\frac{n_1}{2}]+1)}$	$2(n_2-1)$	$2(n_2-1)$, η_{n_1+1} has multiplicity n_2-1.

At the end of this paper, we claim that the only possibility to obtain explicit MLEs is to put constraints on elements of $\boldsymbol{\Sigma}$ and consider a constraint model. The choice of different constraints should be considered in detail, for example, these constraints should not violate the invariance assumption.

5.3 Paper III: On estimation in hierarchical models with block circular covariance structures

Paper III concerns the explicit MLEs of the (co)variance parameters in model (1.2) with a block circular covariance structure, which is the natural continuation of Paper II. As noted from (5.4) the model has three (co)variance parameters more than distinct eigenvalues of the covariance matrix $\boldsymbol{\Sigma}$, and we have to put at least three restrictions on the parameter space to estimate $\boldsymbol{\theta}$ uniquely (and in this case explicitly). Besides guaranteeing the identifiability of the (co)variance parameters, the challenge we face is to preserve the mixed block structure of $\boldsymbol{\Sigma}$ when constraining some of the parameters, which is the main concern of Paper III.

We refer again to Theorem 3.1.2, when the set of covariance parameters can be parameterized by a linear function of canonical parameters, and the number of $\boldsymbol{\theta}$ equals the number of distinct eigenvalues $\boldsymbol{\eta}$ in $\boldsymbol{\Sigma}$, the MLE for $\boldsymbol{\theta}$ has an explicit expression, which is obtained by solving the linear system $\boldsymbol{\eta} = \boldsymbol{L}\boldsymbol{\theta}$, where $\boldsymbol{L}$ is a non-singular coefficient matrix representing how $\boldsymbol{\eta}$ can be expressed by $\boldsymbol{\theta}$.

Theorem 5.3.1 *(Theorem 1, Paper III) Let $\boldsymbol{\eta}$ be a vector of the $2r$ distinct* *eigenvalues of* $\boldsymbol{\Sigma}$ *defined in(1.3). Then* $\boldsymbol{\eta}$ *can be expressed as:*

$$\boldsymbol{\eta} \quad = \quad \boldsymbol{L}\boldsymbol{\theta},$$

where

$$\boldsymbol{L} = (\boldsymbol{B}_1 \vdots \boldsymbol{B}_2),$$

and

$$\boldsymbol{B}_1 = \begin{pmatrix} 1 & n_1 & n_1(n_2-1) \\ \mathbf{1}_{r-1} & \mathbf{0}_{r-1} & \mathbf{0}_{r-1} \\ 1 & n_1 & -n_1 \\ \mathbf{1}_{r-1} & \mathbf{0}_{r-1} & \mathbf{0}_{r-1} \end{pmatrix}, \quad \boldsymbol{B}_2 = \begin{pmatrix} \boldsymbol{A} & (n_2-1)\boldsymbol{A} \\ \boldsymbol{A} & -\boldsymbol{A} \end{pmatrix},$$

$\mathbf{0}_{r-1}$ *is a column vector of size* $r-1$ *with all elements equal to zero, and* $\boldsymbol{A} = (a_{ij})$ *is a square matrix of size* r *with*

$$a_{ij} = \begin{cases} 2^{I(1<j<r)}\cos(2\pi(i-1)(n_1-j+1)/n_1), & \text{if } n_1 \text{ is even}, \\ 2^{I(1<j\le r)}\cos(2\pi(i-1)(n_1-j+1)/n_1), & \text{if } n_1 \text{ is odd}. \end{cases} \tag{5.5}$$

where $I(\cdot)$ *is the indicator function and* $i, j = 1, \ldots, r$.

It is worth observing that the matrix $\boldsymbol{L} : (2r+3) \times 2r$ is of rank $r(\boldsymbol{L}) = 2r$. Our task is to put some restrictions on $\boldsymbol{\theta}$ in (5.5), i.e. $\boldsymbol{K\theta} = \mathbf{0}$, which is equivalent to $\boldsymbol{\theta} = (\boldsymbol{K}')^o\boldsymbol{\theta}^*$, where $(\boldsymbol{K}')^o$: $(2r+3)\times 2r$ is a matrix from which columns generate the orthogonal complement to the column vector space of $\boldsymbol{K}'$ and $\boldsymbol{\theta}_i^*$: $2r \times 1$ is the vector of unknown covariance parameters in model (1.2). Hence, $\boldsymbol{\eta} = \boldsymbol{L\theta} = \boldsymbol{L}(\boldsymbol{K}')^o\boldsymbol{\theta}^*$. If $\boldsymbol{L}(\boldsymbol{K}')^o$ is invertible then $\boldsymbol{\theta}^*$ can be estimated, which can be obtained by $\hat{\boldsymbol{\theta}}^* = (\boldsymbol{L}(\boldsymbol{K}')^o)^{-1}\hat{\boldsymbol{\eta}}$.

We utilize the fact that $\boldsymbol{\eta}$ in (5.5) is not only a function of unknown covariance parameters in $\boldsymbol{\theta}$, $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\theta})$, but also a function of the distinct eigenvalues

$$\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_1} = (\lambda_1^{\boldsymbol{\Sigma}_1}, \lambda_2^{\boldsymbol{\Sigma}_1}), \boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2} = (\lambda_{11}^{\boldsymbol{\Sigma}_2}, \ldots, \lambda_{1r}^{\boldsymbol{\Sigma}_2}, \lambda_{21}^{\boldsymbol{\Sigma}_2}, \ldots, \lambda_{2r}^{\boldsymbol{\Sigma}_2}), \text{ and } \lambda^{\boldsymbol{I}}$$

of $\boldsymbol{\Sigma}_1$ in (1.4), $\boldsymbol{\Sigma}_2$ in (1.5) and $\sigma^2\boldsymbol{I}$, respectively, i.e. $\boldsymbol{\eta} = \boldsymbol{\eta}(\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_1}, \boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2}, \lambda^{\boldsymbol{I}})$:

$$\eta_i = \lambda^{\boldsymbol{I}} + n_1\lambda_h^{\boldsymbol{\Sigma}_1} I(i \in \{1, r+1\}) + \lambda_{hj}^{\boldsymbol{\Sigma}_2},$$

where $h = 1 + I(i \ge r+1)$, $j = i - r(h-1)$ and $i = 1, \ldots, 2r$.

It turns out that instead of putting constraints on $\boldsymbol{\theta}$, it is reasonable to impose constraints on the eigenvalues of the covariance matrices of $\boldsymbol{\gamma}_1$ and

$\boldsymbol{\gamma}_2$, i.e. $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$. Hence we have the following model restrictions, which are called Scenario 1 and Scenario 2:

Scenario 1: One constraint is imposed on the spectrum of $\boldsymbol{\Sigma}_1$ and two constraints on the spectrum of $\boldsymbol{\Sigma}_2$. The two possibilities for imposing constraints are given by

(i) $\lambda_g^{\boldsymbol{\Sigma}_1} = 0$, $\lambda_{g1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{h1}^{\boldsymbol{\Sigma}_2} = 0$, $g, h \in \{1, 2\}$, $g \neq h$;

(ii) $\lambda_g^{\boldsymbol{\Sigma}_1} = 0$, $\lambda_{h1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{ij}^{\boldsymbol{\Sigma}_2} = 0$, $g, h, i \in \{1, 2\}$, $g \neq h$, $j \in \{2, \ldots, r\}$.

Scenario 2: Three constraints are imposed on the spectrum of $\boldsymbol{\Sigma}_2$:

(iii) $\lambda_{g1}^{\boldsymbol{\Sigma}_2} = 0$ and $\lambda_{hj}^{\boldsymbol{\Sigma}_2} = 0$, $g = 1, 2$, $h \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

Using the relationship between the eigenvalues $\boldsymbol{\lambda}^{\boldsymbol{\Sigma}_2}$ of $\boldsymbol{\Sigma}_2$ and the elements of $\boldsymbol{\Sigma}_2$ (see Paper II, Corollary 2.6), with the three different conditions in (i)-(iii), the matrix $\boldsymbol{K}_i$ in $\boldsymbol{K}_i\boldsymbol{\theta} = \mathbf{0}$ can be expressed explicitly, where

$$\boldsymbol{K}_1 = \begin{pmatrix} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \end{pmatrix}, \tag{5.6}$$

$$\boldsymbol{K}_2 = \begin{pmatrix} 0 & 1 & (n_2-1) & \mathbf{0}_r & \mathbf{0}_r \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_j & -(1-n_2)^{2-h}\boldsymbol{a}_j \end{pmatrix}, \tag{5.7}$$

$$\boldsymbol{K}_3 = \begin{pmatrix} 0 & 0 & 0 & \boldsymbol{a}_1 & (n_2-1)\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_1 & -\boldsymbol{a}_1 \\ 0 & 0 & 0 & \boldsymbol{a}_j & -(1-n_2)^{2-h}\boldsymbol{a}_j \end{pmatrix}, \tag{5.8}$$

and $\boldsymbol{a}_1$: $r \times 1$ and $\boldsymbol{a}_j$: $r \times 1$ are the corresponding rows of the matrix $\boldsymbol{A}$ defined via (5.5), $h \in \{1, 2\}$ and $j \in \{2, \ldots, r\}$.

In order to have a better understanding of the meaning of the restrictions, their implications on the factors $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ in model (1.2) has been studied.

Scenario 1: (alternative formulation)

(iv) $\mathbf{1}'_{n_2}\boldsymbol{\gamma}_1 = 0$, $\mathbf{1}'_p\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$;

(v) $\mathbf{1}'_{n_2}\boldsymbol{\gamma}_1 = 0$, $(\boldsymbol{v} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v}^{h-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$, $h \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

Scenario 2: (alternative formulation)

(vi) $(\boldsymbol{v}^{g-1} \otimes \mathbf{1}_{n_1})'\boldsymbol{\gamma}_2 = 0$ and $(\boldsymbol{v}^{h-1} \otimes \boldsymbol{v}_j)'\boldsymbol{\gamma}_2 = 0$, $g = 1, 2$, $h \in \{1, 2\}$, $j \in \{2, \ldots, r\}$.

The obtained alternative formulations of Scenarios 1-2 are advantageous since it is clearly seen that when imposing the restrictions on the eigenvalues of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$, the corresponding eigenvectors will specify the constraints to be imposed on the corresponding factors, which are easily interpretable, and the original symmetry assumptions will be preserved. Nahtman (2006) studied the reparametrization constraints in linear models and demonstrated that in the presence of permutation invariance the classical "sum-to-zero" reparametrization condition of a random factor can be expressed through the spectrum of the corresponding covariance matrix.

Finally, we give the following sufficient conditions for obtaining explicit MLEs of $\boldsymbol{\theta}$:

Theorem 5.3.2 *Model* (1.2) *has explicit and unique MLEs for* $\boldsymbol{\theta}$ *if one of the conditions (i)–(iii) given in Scenarios 1-2 holds.*

The MLEs for the vector of the unknown parameters $\boldsymbol{\theta}_i^*$ in model (1.2) under any restriction given by Scenario 1 or Scenario 2 are the following

$$\hat{\boldsymbol{\theta}}_i^* = (\boldsymbol{L}(\boldsymbol{K}_i')^o)^{-1}\hat{\boldsymbol{\eta}}, \tag{5.9}$$

where $\boldsymbol{K}_i$ is given in (5.6)–(5.8). We have also shown that the estimator $\hat{\boldsymbol{\theta}}_i^*$, $i = 1,2,3$, is a linear combination of independent χ^2-distributed random variables.

It is seen from (5.9) that the MLE $\hat{\boldsymbol{\theta}}_i^*$ is a linear function of $\hat{\boldsymbol{\eta}}$, which are independently χ^2 distributed random variables, see Proposition 1, Paper III. Hence, the distribution of each element in $\hat{\boldsymbol{\theta}}_i^*$ will be correspondingly a linear function of independent χ^2 distributions. The distributional property of Chi-squareness can also be utilized when dealing with hypothesis testing problems.

5.4 Paper IV: Testing in multivariate normal models with block circular covariance structures

As mentioned at the beginning (Chapter 1), this thesis aims to study both estimation and hypothesis testing of a block circular Toeplitz covariance structure for multivariate two-level data. In Paper IV, we have focused on the problem of hypothesis testing, including so-called external tests and internal tests.

5.4.1 External tests

Consider

$$\boldsymbol{Y} = (\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n) \sim N_{p,n}((\mathbf{1}_{n_2} \otimes \boldsymbol{\mu})\mathbf{1}_n', \boldsymbol{\Sigma}, \boldsymbol{I}_n), \tag{5.10}$$

where $N_{p,n}((\mathbf{1}_{n_2} \otimes \boldsymbol{\mu})\mathbf{1}_n', \boldsymbol{\Sigma}, \boldsymbol{I}_n)$ denotes the $p \times n$ matrix normal distribution with mean matrix $(\mathbf{1}_{n_2} \otimes \boldsymbol{\mu})\mathbf{1}_n'$ and $p \times p$ covariance matrix between rows $\boldsymbol{\Sigma}$ and n independent columns.

Three specific structures of $\boldsymbol{\Sigma}$, namely, $\boldsymbol{\Sigma}_I$, $\boldsymbol{\Sigma}_{II}$ and $\boldsymbol{\Sigma}_{III}$, were of interest when dealing with hypothesis testing.

(i) $\boldsymbol{\Sigma}_I = \boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}$, where $\boldsymbol{\Sigma}^{(h)} : n_1 \times n_1$ is an unstructured matrix, $h = 1,2$.

(ii) $\boldsymbol{\Sigma}_{II} = \boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}$, where $\boldsymbol{\Sigma}^{(h)}$, $h = 1,2$, is a CT matrix which depends on r parameters, $r = [n_1/2] + 1$, and the symbol $[\bullet]$ stands for the integer part. For simplicity, the CT matrices can be written as

$$\boldsymbol{\Sigma}^{(1)} = Toep(\tau_1, \tau_2, \tau_3, \ldots, \tau_2),$$
$$\boldsymbol{\Sigma}^{(2)} = Toep(\tau_{r+1}, \tau_{r+2}, \tau_{r+3}, \ldots, \tau_{r+2}).$$

(iii) $\boldsymbol{\Sigma}_{III} = \boldsymbol{I}_{n_2} \otimes \boldsymbol{\Sigma}^{(1)} + (\boldsymbol{J}_{n_2} - \boldsymbol{I}_{n_2}) \otimes \boldsymbol{\Sigma}^{(2)}$, where $\boldsymbol{\Sigma}^{(h)}$, $h = 1,2$, is a CS matrix and can be written as $\boldsymbol{\Sigma}^{(h)} = \sigma_{h1}\boldsymbol{I}_{n_1} + \sigma_{h2}(\boldsymbol{J}_{n_1} - \boldsymbol{I}_{n_1})$.

Observing that the number of unknown parameters in $\boldsymbol{\Sigma}_I$, $\boldsymbol{\Sigma}_{II}$ and $\boldsymbol{\Sigma}_{III}$ are $n_1(n_1+1)$, $2r$ and 4, respectively. We are first interested testing both mean $\boldsymbol{\mu}$ and a block structure of $\boldsymbol{\Sigma}$ simultaneously.

$$H_1^0 : \boldsymbol{\mu} = \mu\mathbf{1}_{n_1}, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II} \quad \text{versus} \quad H_1^a : \boldsymbol{\mu} \in \mathbb{R}^{n_1}, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_I,$$
$$H_2^0 : \boldsymbol{\mu} = \mu\mathbf{1}_{n_1}, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III} \quad \text{versus} \quad H_2^a : \boldsymbol{\mu} \in \mathbb{R}^{n_1}, \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_I.$$

Furthermore, the following hypotheses about patterned covariance matrices, i.e.,

$$H_3^0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III} \quad \text{versus} \quad H_3^a : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II}, \text{ given } \boldsymbol{\mu} = \mu\mathbf{1}_{n_1},$$
$$H_4^0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II} \quad \text{versus} \quad H_4^a : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_I, \text{ given } \boldsymbol{\mu} \in \mathbb{R}^{n_1},$$
$$H_5^0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III} \quad \text{versus} \quad H_5^a : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_I, \text{ given } \boldsymbol{\mu} \in \mathbb{R}^{n_1},$$

or patterned means, i.e.,

$$H_6^0 : \boldsymbol{\mu} = \mu\mathbf{1}_{n_1} \quad \text{versus} \quad H_6^a : \boldsymbol{\mu} \in \mathbb{R}^{n_1}, \text{ given } \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{II},$$
$$H_7^0 : \boldsymbol{\mu} = \mu\mathbf{1}_{n_1} \quad \text{versus} \quad H_7^a : \boldsymbol{\mu} \in \mathbb{R}^{n_1}, \text{ given } \boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{III},$$

can be of interest in various applications.

The LRT statistics Λ_i for testing the null hypotheses H_i^0 versus the alternative hypotheses H_i^a, $i = 1, \ldots, 7$, have been derived. We have also shown that under the null hypothesis H_i^0, the distribution of Λ_i has the same distribution as a product of independent Beta random variables. Here we only present the results for the first testing situation; the other results of the external test are referred to in Paper IV.

Theorem 5.4.1 *(Theorem 3.1, Paper IV) Let* $\boldsymbol{Y}$ *be defined in* (5.10) *and* $\boldsymbol{K}$ *in* (3.5). *Put*

$$\boldsymbol{X}_1 = (n_2^{-1/2}\mathbf{1}'_{n_2} \otimes \boldsymbol{I}_{n_1})\boldsymbol{Y}, \quad \boldsymbol{X}_2 = (\boldsymbol{K}'_1 \otimes \boldsymbol{I}_{n_1})\boldsymbol{Y}, \quad \boldsymbol{X}'_2 = (\boldsymbol{X}'_{21}, \ldots, \boldsymbol{X}'_{2(n_2-1)}).$$

The LRT statistic for testing H_1^0 versus H_1^a is given by

$$\Lambda_1^{2/n} = \frac{2^{n_2(n_1-1-\mathbb{1}_{\text{even}})}|\boldsymbol{S}_1||\boldsymbol{S}_2|^{n_2-1}}{t_{21}\prod_{i=2}^{r} t_{1i}^{m_i}\prod_{i=r+1}^{2r} t_{2i}^{m_i}}, \tag{5.11}$$

where $\mathbb{1}_{\text{even}}$ is the indicator function that n_1 is even, $\boldsymbol{S}_1 = \boldsymbol{X}_1\boldsymbol{Q}_n\boldsymbol{X}_1'$, $\boldsymbol{S}_2 = \sum_{i=1}^{n_2-1}\boldsymbol{X}_{2i}\boldsymbol{X}_{2i}'$, $t_{1i} = \text{tr}((\boldsymbol{X}_1\boldsymbol{X}_1')(\boldsymbol{v}_i\boldsymbol{v}_i' + \boldsymbol{v}_{n_1-i+2}\boldsymbol{v}_{n_1-i+2}'))$, $t_{21} = \text{tr}(\boldsymbol{S}_1\boldsymbol{P}_{n_1})$, $t_{2,r+i} = \text{tr}(\boldsymbol{S}_2(\boldsymbol{v}_i\boldsymbol{v}_i' + \boldsymbol{v}_{n_1-i+2}\boldsymbol{v}_{n_1-i+2}'))$, $i = 2,\dots,r-1$, and for $i = r$,

$$t_{1r} = \begin{cases} \text{tr}((\boldsymbol{X}_1\boldsymbol{X}_1')(\boldsymbol{v}_r\boldsymbol{v}_r' + \boldsymbol{v}_{n_1-r+2}\boldsymbol{v}_{n_1-r+2}')), & \textit{if } n_1 \textit{ is odd,}\\ \text{tr}((\boldsymbol{X}_1\boldsymbol{X}_1')(\boldsymbol{v}_r\boldsymbol{v}_r')), & \textit{if } n_1 \textit{ is even,}\end{cases}$$

$$t_{2,2r} = \begin{cases} \text{tr}(\boldsymbol{S}_2(\boldsymbol{v}_r\boldsymbol{v}_r' + \boldsymbol{v}_{n_1-r+2}\boldsymbol{v}_{n_1-r+2}')), & \textit{if } n_1 \textit{ is odd,}\\ \text{tr}(\boldsymbol{S}_2\boldsymbol{v}_r\boldsymbol{v}_r'), & \textit{if } n_1 \textit{ is even.}\end{cases}$$

Theorem 5.4.2 *(Theorem 3.4, Paper IV) Under the null hypothesis H_1^0, the distribution of Λ_1, given in* (5.11)*, follows*

$$\Lambda_1^{2/n} \overset{d}{\sim} \prod_{i=1}^{n_1-1} B_{1i}B_{2i}^{n_2-1}, \tag{5.12}$$

where B_{1i} and B_{2i} are independent distributed, $i = 1,\dots,n_1-1$,

$$B_{1i} \sim \begin{cases} \beta(\frac{n-i-1}{2}, \frac{i+1}{2}), & \textit{for } i = 1,\dots,[n_1/2],\\ \beta(\frac{n-i-1}{2}, \frac{i+2}{2}), & \textit{for } i = [n_1/2]+1,\dots,n_1-1,\end{cases}$$

$$B_{2i} \sim \begin{cases} \beta(\frac{n(n_2-1)-i}{2}, \frac{i}{2}), & \textit{for } i = 1,\dots,[n_1/2],\\ \beta(\frac{n(n_2-1)-i}{2}, \frac{i+1}{2}), & \textit{for } i = [n_1/2]+1,\dots,n_1-1.\end{cases}$$

Based on the null distribution of each test statistic, the asymptotic expansion for the distribution function of LRT statistic as given by Box (1949) can be given.

5.4.2 Internal test

In this section, we consider testing hypotheses about (co)variance parameters of $\boldsymbol{\Sigma}$ in model (1.2).

Our main concern is devoted to the discussion of the testability concerning the (co)variance parameters of $\boldsymbol{V}(\boldsymbol{\theta})$ in model (1.2). Here we change the notation in order to distinguish with $\boldsymbol{\Sigma}$ in model (5.10). The testability of the fixed effects in the framework of the linear model has been investigated see Roy and Roy (1959) and Das Gupta (1977), for example. However, the testability of the random effects is rarely paid attention to and has even not been well defined. Here, the testability is only restricted to the "possibility" of testing (co)variance parameters, which can be done through the linear restrictions of the eigenvalues of $\boldsymbol{V}(\boldsymbol{\theta})$. Our underlying idea is to make use of the knowledge of the eigenvalues of $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{\eta}$ in Table 5.1, when developing the test procedures concerning the three restricted models $\mathcal{M}_i$ given by Paper III, $i = 1, 2, 3$.

We consider the hypothesis testing problems of $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ separately for each restricted model $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$. In model $\mathcal{M}_1$ and $\mathcal{M}_2$ we will test

$$H_{01} : \sigma_2 = 0 \quad \text{versus} \quad H_{a1} : \sigma_2 < 0,$$

and in $\mathcal{M}_3$ we will test

$$H_{02} : \sigma_2 = 0 \quad \text{versus} \quad H_{a2} : -\frac{\sigma_1}{n_2 - 1} \leq \sigma_2 \leq \sigma_1.$$

The hypothesis H_{01} implies that there is no random effect $\boldsymbol{\gamma}_1$ in model $\mathcal{M}_1$ and $\mathcal{M}_2$, while the hypothesis H_{02} means that the factor levels of $\boldsymbol{\gamma}_2$ are uncorrelated.

Then, testing H_{01} versus H_{a1} in model $\mathcal{M}_1$ is equivalent to testing

$$H_{01,\mathcal{M}_1} : \eta_1 = \eta_{r+1} \text{ versus } H_{a1,\mathcal{M}_1} : \eta_1 < \eta_{r+1}, \tag{5.13}$$

and testing H_{01} versus H_{a1} in model $\mathcal{M}_2$ is equivalent to testing

$$H_{01,\mathcal{M}_2} : \eta_l = \eta_{r+1} \text{ versus } H_{a1,\mathcal{M}_2} : \eta_l < \eta_{r+1}, \quad \text{for some } l \in \{2, \ldots, r, r+2, \ldots, 2r\}. \tag{5.14}$$

In model $\mathcal{M}_3$, testing H_{02} versus H_{a2} is equivalent to testing

$$H_{02} : \eta_1 = \eta_{r+1} \text{ versus } H_{a2} : \eta_1 \neq \eta_{r+1}. \tag{5.15}$$

Concerning the covariance matrix $\boldsymbol{\Sigma}_2$ in (1.5), we will test for $\mathcal{M}_1$, $\mathcal{M}_2$ and $\mathcal{M}_3$

$$H_{03} : \tau_1 = 0 \quad \text{versus} \quad H_{a3} : \tau_1 \neq 0.$$

The hypothesis H_{03} means that there is no random effect $\boldsymbol{\gamma}_2$ in all restricted models, and it implies that testing H_{03} versus H_{a3} in the restricted models $\mathcal{M}_1$ and $\mathcal{M}_2$ is equivalent to testing

$$H_{03,\mathcal{M}_i} : \eta_l = \sigma^2 \text{ versus } H_{a3,\mathcal{M}_i} : \text{ none of } \eta_l \text{ is equal, } i \in \{1, 2\}, l \neq r+1, \tag{5.16}$$

and in model $\mathcal{M}_3$, the testing problem can be formulated as

$$H_{03,\mathcal{M}_3} : \eta_l = \sigma^2 \text{ versus } H_{a3,\mathcal{M}_3} : \text{ none of } \eta_l \text{ is equal, } \quad l \neq 1, r+1. \tag{5.17}$$

It can be observed that the hypotheses H_{01} and H_{03} are on the boundaries of the parameter space, since we have $\sigma_1 = -(n_2-1)\sigma_2$, where σ_1 is the variance of $\mathbf{\Sigma}_1$, and τ_1 is the variance of $\mathbf{\Sigma}_2$. The derived results above are essential since it shows that the original parameter spaces of the (co)variance parameters do not need to be taken into account, and it indicates that testing the hypotheses of H_{01}, H_{02} and H_{03} is nothing but testing the equality of several distinct eigenvalues which are in fact the variances in the context of the canonical form of model (1.2), see (5.13)-(5.17). Hence we can construct Bartlett's test statistics (or other alternatives for testing the equality of variances) and the results are given in Section 4, Paper IV.

6. Concluding remarks, discussion and future research

6.1 Contributions of the thesis

The main contribution of this thesis is that a new block covariance structure for characterizing the dependency of specific multivariate two-level data has been introduced, and inferential results including both explicit MLEs and hypothesis testing have been obtained. The contributions of each paper are summarized as follows:

- In Paper I, two patterned covariance matrices are derived under specific invariance assumption. The derivation extends the specification of covariance structures within the group invariance framework and the spectra of the two types of patterned covariance matrices are obtained.

- A random effects model with a block circular Toeplitz covariance structure is considered in Paper II. The spectral properties of the corresponding covariance matrix are derived and the canonical parameters can be expressed as linear functions of (co)variance parameters. It is proved that explicit MLEs of (co)variance parameters do not exist without imposing constraints on the parameter space.

- The discussion of explicit estimation in model (1.2) with block circular Toeplitz covariance structures is continued in Paper III. Sufficient conditions for obtaining explicit estimators of (co)variance components are derived. The advantage of this approach is that the corresponding eigenvectors will specify the constraints to be imposed on the factor, which are usually interpretable, at the same time keeping the original symmetry assumptions.

- Hypothesis testing for the nested multivariate normal models with block circular covariance structures is discussed in Paper IV. Tests concerning the general block structures of the covariance matrix (external tests) and tests for specific (co)variance components (internal tests) are constructed. Likelihood ratio test (LRT) statistics and the corresponding null distributions are obtained.

6.2 Discussion

The aims of this thesis have been considered and discussed across relevant papers.

Szatrowski's result of explicit MLEs of (co)variance parameters is conditioned at the existence of the explicit MLE of the mean parameters, which indicates that for a classical mixed linear model in (3.1), the column space of $\boldsymbol{X}$ must be $\boldsymbol{\Sigma}$-invariant. This is not a trivial condition for any mean structure, for example, in the growth curve model. The general bilinear mean structure $\boldsymbol{ABC}$, where $\boldsymbol{A}$, $\boldsymbol{C}$ are known matrices and $\boldsymbol{B}$ is a matrix containing the unknown mean parameters, does not preserve the $\boldsymbol{\Sigma}$-invariant property unless some conditions are imposed on the matrix $\boldsymbol{A}$.

Choosing appropriate restrictions on the parameters $\boldsymbol{\theta}$ for model (1.2) may sometimes be a difficult task as discussed in Paper III. One reason is that the existence of an explicit estimator for the mean parameter (μ) should not be affected in the restricted models. Imposing restrictions on the spectra of the covariance matrices is beneficial from a $\boldsymbol{\Sigma}$-invariant point of view, since in this case we can find the restrictions that can preserve the block structure of $\boldsymbol{\Sigma}$, at the same time coming into the existence of the explicit estimator of μ. Despite the positive definiteness of $\boldsymbol{\Sigma}$ being guaranteed, we still face the problem of possible negative definiteness for $\boldsymbol{\Sigma}_1$ or $\boldsymbol{\Sigma}_2$ since the restricted models have not taken the issue of non negative definiteness into account. It is also worth noting that only the observed data and the problem at hand can guide us concerning which model to choose since the maximum of the likelihood function of model (1.2) will not be affected by the choice of the restricted models.

The idea with the internal test is new. Investigating the testability of the parameters in the block circular covariance matrix yields that testing some parameters can be done by testing the equality of several eigenvalues of a patterned covariance matrix, which is advantageous since the theory of testing equal variances has been well developed. However, it is not obvious that one can find a testable hypothesis in this way and it replies on the knowledge of eigenvalues.

It can be observed that there is some connection between the external and internal test since in the internal test, the equality of some eigenvalues for the block circular covariance structure implies that the matrix will "degenerate" to a more parsimonious pattern, for example, $\boldsymbol{\Sigma}_{BCS-CS}$ in (2.13), and the hypothesis becomes exactly the same as one of the hypotheses (H_3^0)

in the external one.

6.3 Future research

Model (1.2) can be extended by considering other types of mean structures, e.g. profiles in profile analysis. Srivastava and Singull (2012) have studied profile analysis when the random effect has a CS covariance structure, and their study can be extended when the random effect has a block circular Toeplitz covariance structure.

In many medical studies, it is necessary to investigate the effect to patients after receiving a particular therapy. Recently, Roy *et al.* (2015) developed a test statistic for testing the equality of mean vectors under a block compound symmetry (BCS) covariance structure in (2.9), which is an extension of the Hotelling's T^2 statistic. Testing the equality of mean structures when observations exhibit the block circular Toeplitz covariance structures has not been explored yet and this gap can possibly be filled.

We can relax the assumption of n independent samples in (1.2). Instead of $\boldsymbol{Y} \sim N_{p,n}(\mu \mathbf{1}_p \mathbf{1}_n', \boldsymbol{\Sigma}, \boldsymbol{I}_n)$, we assume $\boldsymbol{Y} \sim N_{p,n}(\mu \mathbf{1}_p \mathbf{1}_n', \boldsymbol{\Sigma}, \boldsymbol{\Psi})$, where $\boldsymbol{\Psi}$ is an $n \times n$ covariance matrix between n columns. Besides $\boldsymbol{\Sigma}$ having a block circular Toeplitz matrix, some structures can also be imposed to $\boldsymbol{\Psi}$.

One possible extension of Paper IV could strive to test a more parsimonious pattern than $\boldsymbol{\Sigma}_{BCS-CS}$ given in (2.13). Recall that the covariance structure $\boldsymbol{\Sigma}_{DCS}$ in (2.10) has three unknown parameters, and it is a more parsimonious pattern than $\boldsymbol{\Sigma}_{BCS-CS}$ which contains four parameters. The external test can be incorporated with more hypotheses such as testing $\boldsymbol{\Sigma}_{DCS}$ versus $\boldsymbol{\Sigma}_{BCS-CS}$ or $\boldsymbol{\Sigma}_{DCS}$ versus the block circular Toeplitz structure in (1.3).

7. Sammanfattning

I denna avhandling studeras både skattning-s och hypotesprövningsproblem i balanserade multivariata modeller med slumpmässiga effekter med en specifik kovariansstruktur, som kallas blockcirkulär Toeplitz. Denna kovariansstruktur beskriver beroendet i data med två nivåer och olika symmetriegenskaper.

Vi härleder två kovariansstrukturer under två olika invariansrestriktioner (symmetri). De erhållna kovariansstrukturerna speglar både cirkuläritet och utbytbarhet. Nya uttryck för egenvärdena av blockcirkulära symmetriska matriser erhålls som tar hänsyn till blockstrukturerna.

Skattning av parametrarna i de balanserade slumpmässiga effektmodellerna med blockcirkulära kovariansmatriser betraktas. De spektrala egenskaperna hos sådana kovariansstrukturer tillhandahålls. Vi härleder maximum-likelihoodskattningen genom spektralspjälkning av kovariansmatrisen och diskuterar förekomsten av explicita maximum likelihoodskattningar för kovariansparametrarna. Tillräckliga villkor för att erhålla explicita och unika skattningar för varians-kovarianskomponenter har härletts. Olika begränsade modeller diskuteras och motsvarande maximum likelihoodskattningar presenteras.

Avhandlingen behandlar också hypotesprövning av blockkovariansstrukturer, speciellt blockcirkulära Toeplitzkovariansmatriser. Vi studerar både så kallade externa och interna tester. I de externa testerna tar vi upp olika hypotestest kring blockkovariansstrukturerna samt medelvärdsstrukturer. De interna testerna handlar om att testa specifika kovariansparametrar i de blockcirkulära Toeplitzmatriserna. Likelihood-kvottest konstrueras och fördelningar av motsvarande teststatistikor härleds under nollhypoteserna.

References

Anderson, T. W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. *The Annals of Statistics,* **1**, 135–141. 2, 9, 34

Anderson, T. W. (2003). *An introduction to multivariate statistical analysis, 3rd ed.* Wiley, New Jersey. 27, 28, 29

Andersson, S. (1975). Invariant normal models. *The Annals of Statistics,* **3**, 132–154. 12

Andersson, S. and Madsen, J. (1998). Symmetry and lattice conditional independence in a multivariate normal distribution. *The Annals of Statistics,* **26**, 525–572. 12

Arnold, S. F. (1973). Application of the theory of products of problems to certain patterned covariance matrices. *The Annals of Statistics,* **1**, 682–699. 2, 14, 22

Arnold, S. F. (1979). Linear models with exchangeably distributed errors. *Journal of the American Statistical Association,* **74**, 194–199. 3, 14, 23

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences,* **160**, 268–282. 30

Barton, T. A. and Fuhrmann, D. R. (1993). Covariance structures for multidimensional data. *Multidimensional Systems and Signal Processing,* **4**, 111–123. 32

Basilevsky, A. (1983). *Applied matrix algebra in the statistical sciences.* North-Holland, New York. 22

Box, G. E. P. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika,* **36**, 317–346. 27, 28, 40

Chi, E. M. and Reinsel, G. C. (1989). Models for longitudinal data with random effects and ar (1) errors. *Journal of the American Statistical Association,* **84**, 452–459. 28

Crainiceanu, C. M. and Ruppert, D. (2004). Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **66**, 165–185. 30

Das Gupta, S. (1977). Two problems in multivariate analysis: BLUS residuals and testability of linear hypothesis. *Annals of the Institute of Statistical Mathematics*, **29**, 35–41. 40

Davidov, E. and Depner, F. (2011). Testing for measurement equivalence of human values across online and paper-and-pencil surveys. *Quality & Quantity*, **45**, 375–390. 5

Demidenko, E. (2004). *Mixed models: theory and applications*. Wiley, New York. 19

Eaton, M. L. (1983). *Multivariate statistics: a vector space approach*. Wiley, New York. 12, 13

Fitzmaurice, G. M., Laird, N. M. and Ware, J. H. (2004). *Applied longitudinal analysis*. Wiley, Hoboken, New Jersey. 1

Fleiss, J. L. (1966). Assessing the accuracy of multivariate observations. *Journal of the American Statistical Association*, **61**, 403–412. 2

Gallo, J. and Khuri, A. I. (1990). Exact tests for the random and fixed effects in an unbalanced mixed two-way cross-classification model. *Biometrics*, **46**, 1087–1095. 29

Geisser, S. (1963). Multivariate analysis of variance for a special covariance case. *Journal of the American Statistical Association*, **58**, 660–669. 2

Gurtman, M. (2010). Circular reasoning about circular assessment. In L. Horowitz and S. Strack, eds., *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions*. John Wiley & Sons, New Jersey. 5

Jennrich, R. I. and Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, **42**, 805–820. 22

Khatri, C. G. (1973). Testing some covariance structures under a growth curve model. *Journal of Multivariate Analysis*, **3**, 102–116. 2

Kollo, T. and von Rosen, D. (2005). *Advanced multivariate statistics with matrices*. Springer-Verlag, Dordrecht. 21, 32

Krishnaiah, P. and Lee, J. C. (1974). On covariance structures. *Sankhyā: The Indian Journal of Statistics, Series A,* , 357–371. 2, 22

Krishnaiah, P. R. and Lee, J. C. (1980). Likelihood ratio tests for mean vectors and covariance matrices. *Handbook of Statistics,* **1**, 513–570. 27

Louden, C. and Roy, A. (2010). Classification rules under autoregressive and general circulant covariance. *Communications in Statistics–Theory and Methods,* **39**, 3294–3315. 5

Marden, J. I. (2012). "Multivariate statistics old school". Department of Statistics, University of Illinois at Urbana-Champaign. 12

Marin, J. M. and Dhorne, T. (2002). Linear Toeplitz covariance structure models with optimal estimators of variance components. *Linear Algebra and its Applications,* **354**, 195–212. 10, 11

Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics,* **11**, 204–209. 12

Mitra, S. and Moore, B. (1973). Gauss-Markov estimation with an incorrect dispersion matrix. *Sankhyā: The Indian Journal of Statistics. Series A,* **35**, 139–152. 20

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory.* Wiley, New York. 29

Nahtman, T. (2006). Marginal permutation invariant covariance matrices with applications to linear models. *Linear Algebra and its Applications,* **417**, 183–210. 12, 13, 16, 23, 37

Nahtman, T. and von Rosen, D. (2008). Shift permutation invariance in linear random factor models. *Mathematical Methods of Statistics,* **17**, 173–185. 12, 13, 16, 17, 24

Öfversten, J. (1993). Exact tests for variance components in unbalanced mixed linear models. *Biometrics,* **49**, 45–57. 29

Ohlson, M., Andrushchenko, Z. and von Rosen, D. (2011). Explicit estimators under m-dependence for a multivariate normal distribution. *Annals of the Institute of Statistical Mathematics,* **63**, 29–42. 11

Olkin, I. (1973a). Inference for a normal population when the parameters exhibit some structure. Tech. Rep. No.77, Department of Statistics, Stanford University. 2, 14

Olkin, I. (1973b). Testing and estimation for structures which are circularly symmetric in blocks. In D. G. Kabe and R. P. Gupta, eds., *Multivariate statistical inference*. 183-195, North–Holland, Amsterdam. 2, 12, 15, 16, 24, 28, 29, 33

Olkin, I. and Press, S. (1969). Testing and estimation for a circular stationary model. *The Annals of Mathematical Statistics*, **40**, 1358–1373. 4, 12, 16, 22, 28, 29, 33

Pan, J. X. and Fang, K. T. (2002). *Growth curve model and statistical diagnostics*. Springer, New York. 21

Perlman, M. D. (1987). Group symmetry covariance models. *Statistical Science*, **2**, 421–425. 12

Puntanen, S. and Styan, G. P. (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator. *The American Statistician*, **43**, 153–161. 20, 21

Rao, C. R. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1. 355–372, Univ. of California Press Berkeley. 21

Roebruck, P. (1982). Canonical forms and tests of hypotheses: Part I: The general univariate mixed model. *Statistica Neerlandica*, **36**, 63–74. 21

von Rosen, T. (2011). On the inverse of certain block structured matrices generated by linear combinations of Kronecker products. *Linear and Multilinear Algebra*, **59**, 595–606. 12

Roy, A. and Fonseca, M. (2012). Linear models with doubly exchangeable distributed errors. *Communications in Statistics–Theory and Methods*, **41**, 2545–2569. 3, 15, 24

Roy, A. and Leiva, R. (2008). Likelihood ratio tests for triply multivariate data with structured correlation on spatial repeated measurements. *Statistics & Probability Letters*, **78**, 1971–1980. 3

Roy, A. and Leiva, R. (2011). Estimating and testing a structured covariance matrix for three-level multivariate data. *Communications in Statistics-Theory and Methods*, **40**, 1945–1963. 14

Roy, A., Leiva, R., Žežula, I. and Klein, D. (2015). Testing the equality of mean vectors for paired doubly multivariate observations in blocked compound symmetric covariance matrix setup. *Journal of Multivariate Analysis*, **137**, 50–60. 45

Roy, S. and Roy, J. (1959). A note on a class of problems in 'normal' multivariate analysis of variance. *The Annals of Mathematical Statistics*, **30**, 577–581. 40

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna, ed., *Advances in experimental social psychology*, vol. 25. 1–65, Academic Press, New York. 5

Seely, J. F. and El-Bassiouni, Y. (1983). Applying Wald's variance component test. *Annals of Statistics*, **11**, 197–201. 29

Self, S. G. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, **82**, 605–610. 29

Shi, J. H. and Wang, S. G. (2006). The spectral decomposition of covariance matrices for the variance components models. *Journal of Multivariate Analysis*, **97**, 2190–2205. 21

Srivastava, M. and Singull, M. (2012). Profile analysis with random-effects covariance structure. *Journal of the Japan Statistical Society*, **42**, 145–164. 30, 45

Srivastava, M. S., von Rosen, T. and von Rosen, D. (2008). Models with a Kronecker product covariance structure: estimation and testing. *Mathematical Methods of Statistics*, **17**, 357–370. 17

Srivastava, M. S., von Rosen, T. and von Rosen, D. (2009). Estimation and testing in general multivariate linear models with Kronecker product covariance structure. *Sankhyā: The Indian Journal of Statistics, Series A*, , 137–163. 17

Steinmetz, H., Isidor, R. and Baeuerle, N. (2012). Testing the circular structure of human values: A meta-analytical structural equation modelling approach. *Survey Research Methods*, **6**, 61–75. 5

Szatrowski, T. H. (1980). Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. *The Annals of Statistics*, **8**, 802–810. 19, 20, 21, 34

Szatrowski, T. H. and Miller, J. J. (1980). Explicit maximum likelihood estimates from balanced data in the mixed model of the analysis of variance. *The Annals of Statistics*, **8**, 811–819. 22

Viana, M. and Olkin, I. (2000). Symmetrically dependent models arising in visual assessment data. *Biometrics*, **56**, 1188–1191. 3

Votaw, D. F. (1948). Testing compound symmetry in a normal multivariate distribution. *The Annals of Mathematical Statistics*, **19**, 447–473. 2, 12, 28

Wald, A. (1941). On the analysis of variance in case of multiple classifications with unequal class frequencies. *The Annals of Mathematical Statistics*, **12**, 346–350. 29

Wald, A. (1947). A note on regression analysis. *The Annals of Mathematical Statistics*, **18**, 586–589. 29

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, **9**, 60–62. 28

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**, 257–281. 2, 12, 28

Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models. *The Annals of Mathematical Statistics*, **38**, 1092–1109. 20