Analysis of Cross-Over Designs via Growth Curve Models

Nils Åsenblad and Dietrich von Rosen

Research Report
Centre of Biostochastics

Swedish University of Agricultural Sciences

Report 2004:3
ISSN 1651-8543

Analysis of Cross-Over Designs via Growth Curve Models

Nils Åsenblad and Dietrich von Rosen

Centre of Biostochastics
Swedish University of Agricultural Sciences

Abstract

Models for repeated measures cross-over designs are defined in terms of growth curve models. In the paper two specific cross-over designs, called the AB:BA and ABAB:BABA design, are studied. The maximum likelihood estimators for the parameters are derived by utilizing the theory for growth curve models. A model with a structured dispersion matrix is defined for the AB:BA design, and a specific linear transformation is used to derive estimators in a convenient way. To illustrate numerically, maximum likelihood estimates are calculated for an ABAB:BABA study.

Keywords: Cross-over design, growth curve models, repeated measurements, maximum likelihood estimator.

AMS 1991 subject classifications: Primary 62H12; Secondary 62K99

1 Introduction

The simplest example of a cross-over design is the two-sequence two-period cross-over design for comparing two treatments. That is, if the two treatments involved are denoted A and B, then half of the subjects receive treatment A in the first period followed by treatment B in the second period, whereas the remaining subjects receive the treatments in the reverse order. This design is denoted AB:BA, where the colon is used to separate the treatment sequences. Higher order cross-over designs are obtained by including more than two treatments or two periods. Typically, in a standard cross-over design, a response variable is measured once at the end of each period for each subject. A repeated measures cross-over design is a cross-over design where a sequence of measurements is collected within each period. For example, to compare two test drugs, the concentration of the drug in the blood might be measured every 30 minutes for each subject during three hours after administration of the drug.

The aim of this paper is to show that statistical models for a repeated measures cross-over design in a useful way can be defined through the growth curve model. The growth curve model (GCM) (other names of the model are generalized linear model, GMANOVA, multivariate linear normal model etc.) was introduced by Potthoff and Roy (1964) as a method for analysis of growth curve experiments. An extended version of the GCM was presented by von Rosen (1989). In general, the model applies to longitudinal data in which subjects are followed for a period of time. To date the GCM has only been applied to cross-over designs by Banken (1984), in a canonical version, to study hypothesis testing problems. Putt and Chinchilli (1999) developed a mixed effects model and gave conditions under which explicit estimators of the variance components exist as well as hypothesis test of treatment differences. However, explicit maximum likelihood estimators of the fixed effects are not obtained, unless independent random errors are assumed, in which case the estimators coincide with the ordinary least square estimators. The main advantages of using the GCM is that explicit maximum likelihood estimators of the parameters are obtained.

Two specific cross-over designs, called the AB:BA and ABAB:BABA design, are studied. Models for these designs are described and the maximum likelihood estimators for the parameters are derived by utilizing the theory for the GCM. Two models are defined for each of the design, with respect to inclusion or exclusion of a so-called carry-over effect. When the carry-over effect is omitted, the extended GCM is applied. Furthermore, a model with a structured dispersion matrix is defined for the AB:BA design, and a specific linear transformation is used to derive estimators in a convenient way. To illustrate numerically, maximum likelihood estimates are calculated for an ABAB:BABA study published by Ciminera and Wolfe (1953) and later re-analyzed by Putt and Chinchilli (1999).

2 Background

2.1 Growth curve models

For reviews of the model see von Rosen (1991) and Kshirsagar and Smith (1995). We will define two types of the GCM which will be referred to as GCM1 and GCM2. GCM1 is the same model as introduced by Potthoff and Roy, whereas GCM2 is an extended version of GCM1 introduced by von Rosen (1989). In the following, $\mathbf{X}$ will be a random observable matrix, where each subject or unit is represented by a column in $\mathbf{X}$. Let further, for an arbitrary matrix $\mathbf{A}$ of real numbers, $\rho(\mathbf{A})$ denote the rank of $\mathbf{A}$, $\mathcal{C}(\mathbf{A})$ the column space of $\mathbf{A}$, $\mathbf{A}^-$ a generalized inverse of $\mathbf{A}$, $\mathbf{A}'$ the transpose of $\mathbf{A}$, $\mathcal{C}(\mathbf{A})^\perp$ the orthogonal complement of $\mathcal{C}(\mathbf{A})$ and $\mathbf{A}^o$ any matrix generating $\mathcal{C}(\mathbf{A})^\perp$.

Definition 2.1 (GCM1) *Let $\mathbf{X} : p \times n$, $\mathbf{A} : p \times q$, $q \leq p$, $\mathbf{B} : q \times k$, $\mathbf{C} : k \times n$, $\rho(\mathbf{C}) + p \leq n$ and dispersion matrix $\mathbf{\Sigma} : p \times p$, p.d. The GCM1 for $\mathbf{X}$ is:*

$$\mathbf{X} = \mathbf{ABC} + \mathbf{\Sigma}^{1/2}\mathbf{E},$$

where the elements, e_{kl}, $k = 1, \ldots, p$, $l = 1, \ldots, n$, of $\mathbf{E}$ are iid $\mathbf{N}(0, 1)$, $\mathbf{A}$ and $\mathbf{C}$ are known matrices, and $\mathbf{B}$ and $\mathbf{\Sigma}$ are unknown parameter matrices.

The matrices $\mathbf{A}$ and $\mathbf{C}$ are called the within subject design and between subject design matrices, respectively. Note that $E[\mathbf{X}] = \mathbf{ABC}$ and that the columns of $\mathbf{X}$ are independently distributed according to the multivariate normal distribution $N_p(\mathbf{ABC}_i, \mathbf{\Sigma})$, $i = 1, \ldots, n$. Furthermore, if $\mathbf{A} = \mathbf{I}_p$, where $\mathbf{I}_p$ denotes the identity matrix of size $p \times p$, the ordinary MANOVA model is obtained.

Example 2.1 *Suppose that we study a group of animals under three different treatment conditions and that a time series of p ($p \geq 3$) measurement, $t_1, \ldots, t_p$, of some response variable is collected for each animal. The p-vectors of measurements from animals are assumed to be independent observations from a multivariate normal distribution. Furthermore, suppose that the expected "growth curve" follows a cubic polynomial function in time, where the parameters may differ between treatment groups. Let the first n_1 columns of $\mathbf{X}$*

represent animals in treatment group 1 and the following n_2 columns animals in treatment group 2, and so on. Then

$$
\begin{aligned}
\mathbf{A} &= \begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ 1 & t_2 & t_2^2 & t_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & t_p & t_p^2 & t_p^3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \\ b_{41} & b_{42} & b_{43} \end{pmatrix}, \\
\mathbf{C} &= \begin{pmatrix} \mathbf{1}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{1}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{1}'_{n_3} \end{pmatrix},
\end{aligned}
$$

where n_i is the number of animals in treatment group i, $i = 1, 2, 3$, and $\mathbf{1}_{n_i}$ and $\mathbf{0}_{n_i}$ represent $n_i \times 1$ vectors consisting of ones and zeroes, respectively. The columns of $\mathbf{B}$ represent the unknown parameters for each treatment group, and the expected response at time t equals $b_{1i} + b_{2i}t + b_{3i}t^2 + b_{4i}t^3$ in treatment group i.

Next, an extension of GCM1, called GCM2, will be defined. Note that Definition 2.1 implies that the expected responses within subjects, must follow the structure given in $\mathbf{A}$. For instance, in Example 2.1 all subjects have to follow a cubic polynomial growth. However, the coefficients of the cubic polynomial are allowed to be different between treatment groups. The extended model allow subjects to have a different degree in the polynomial growth curves. In addition, the model is useful when linear restrictions of $\mathbf{B}$ exists in the GCM1, such as $\mathbf{DBF} = \mathbf{0}$, where $\mathbf{D}$ and $\mathbf{F}$ are known matrices.

Definition 2.2 (GCM2) *Let $\mathbf{X} : p \times n$, $\mathbf{A}_i : p \times q_i$, $\mathbf{B}_i : q_i \times k_i$, $\mathbf{C}_i : k_i \times n$, $\rho(\mathbf{C}_i) + p \leq n$, $i = 1, 2, \ldots, m$, $\mathcal{C}(\mathbf{C}'_i) \subseteq \mathcal{C}(\mathbf{C}'_{i-1})$, $i = 2, 3, \ldots, m$ and $\mathbf{\Sigma} : p \times p$, p.d. The GCM2 for $\mathbf{X}$ is:*

$$
\mathbf{X} = \sum_{i=1}^{m} \mathbf{A}_i \mathbf{B}_i \mathbf{C}_i + \mathbf{\Sigma}^{1/2}\mathbf{E},
$$

where the elements, e_{kl}, $k = 1, \ldots, p$, $l = 1, \ldots, n$, of $\mathbf{E}$ are iid $N(0, 1)$, $\mathbf{A}_i$ and $\mathbf{C}_i$ are known matrices and $\mathbf{B}_i$ and $\mathbf{\Sigma}$ are unknown parameter matrices.

It follows that the columns of $\mathbf{X}$ are independently distributed according to the multivariate normal.

Example 2.2 *(Example 2.1 continued). Suppose that the expected response at time t in group 1, 2 and 3 is a linear, quadratic and cubic function of t,* *respectively. That is*

$$\begin{aligned} \textit{expected mean group } 1 &= b_{11} + b_{21}t, \\ \textit{expected mean group } 2 &= b_{12} + b_{22}t + b_{32}t^2, \\ \textit{expected mean group } 3 &= b_{13} + b_{23}t + b_{33}t^2 + b_{43}t^3. \end{aligned}$$

These assumptions can simultaneously be described by the model

$$E[\mathbf{X}] = \mathbf{A}_1\mathbf{B}_1\mathbf{C}_1 + \mathbf{A}_2\mathbf{B}_2\mathbf{C}_2 + \mathbf{A}_3\mathbf{B}_3\mathbf{C}_3,$$

where

$$\mathbf{A}_1 = \begin{pmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_p \end{pmatrix}, \quad \mathbf{A}_2 = \begin{pmatrix} t_1^2 \\ t_2^2 \\ \vdots \\ t_p^2 \end{pmatrix}, \quad \mathbf{A}_3 = \begin{pmatrix} t_1^3 \\ t_2^3 \\ \vdots \\ t_p^3 \end{pmatrix},$$

$$\mathbf{B}_1 = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}, \quad \mathbf{B}_2 = \begin{pmatrix} b_{31} & b_{32} & b_{33} \end{pmatrix}, \quad \mathbf{B}_3 = \begin{pmatrix} b_{41} & b_{42} & b_{43} \end{pmatrix},$$

$$\mathbf{C}_1 = \begin{pmatrix} \mathbf{1}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{1}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{1}'_{n_3} \end{pmatrix}, \quad \mathbf{C}_2 = \begin{pmatrix} \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{1}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{1}'_{n_3} \end{pmatrix}, \quad \mathbf{C}_3 = \begin{pmatrix} \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{0}'_{n_3} \\ \mathbf{0}'_{n_1} & \mathbf{0}'_{n_2} & \mathbf{1}'_{n_3} \end{pmatrix}.$$

Note that the condition $\mathcal{C}(\mathbf{C}'_3) \subseteq \mathcal{C}(\mathbf{C}'_2) \subseteq \mathcal{C}(\mathbf{C}'_1)$ *is fulfilled.*

2.2 Results on growth curve models

2.2.1 Maximum likelihood estimators

The maximum likelihood (ML) estimators of the parameters $\mathbf{B}$ and $\mathbf{\Sigma}$ in the GCM1 are presented in the next lemma. When estimating $\mathbf{B}$ full-rank conditions for $\mathbf{A}$ and $\mathbf{C}$ are supposed to hold. Also, uniqueness conditions for the parameters and given linear combinations of $\widehat{\mathbf{B}}$ are presented as well as the dispersion matrix $D[\widehat{\mathbf{B}}]$. Let $\mathbf{G}^-$ denote an arbitrary g-inverse in the sense of $\mathbf{G}\mathbf{G}^-\mathbf{G} = \mathbf{G}$. Proofs of the lemma as well as results and proofs in the case of general rank conditions can be found in von Rosen (1989, 1990).

Lemma 2.1 (i) *Suppose that* $\rho(\mathbf{A}) = q$ *and* $\rho(\mathbf{C}) = k$ *in the GCM1. The ML estimator of* $\mathbf{B}$ *is given by*

$$\widehat{\mathbf{B}} = (\mathbf{A}'\mathbf{S}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{S}^{-1}\mathbf{X}\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-1}.$$

The ML estimator of $\mathbf{\Sigma}$ *equals*

$$n\widehat{\mathbf{\Sigma}} = (\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C})(\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C})' = \mathbf{S} + \mathbf{V}\mathbf{V}',$$

where $\mathbf{S} = \mathbf{X}(\mathbf{I} - \mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C})\mathbf{X}'$ *and* $\mathbf{V} = \mathbf{X}\mathbf{C}'(\mathbf{C}\mathbf{C}')^{-}\mathbf{C} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C}$.

(ii)

$$D[\widehat{\mathbf{B}}] = c_1(\mathbf{C}\mathbf{C}')^{-1} \otimes (\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A})^{-1},$$

where

$$c_1 = \tfrac{n-k-1}{n-k-p+q-1}.$$

(iii) *An unbiased estimator of* $D[\widehat{\mathbf{B}}]$ *is given by*

$$\widehat{D[\widehat{\mathbf{B}}]} = c_{11}(\mathbf{C}\mathbf{C}')^{-1} \otimes (\mathbf{A}'\widehat{\mathbf{\Sigma}}^{-1}\mathbf{A})^{-1},$$

where

$$c_{11} = \tfrac{n}{n-k-p+q}c_1.$$

(iv) *The ML estimator of* $\mathbf{\Sigma}$ *is always unique. The ML estimator of* $\mathbf{B}$ *is unique if and only if* $\rho(\mathbf{A}) = q$ *and* $\rho(\mathbf{C}) = k$*. The linear combination* $\mathbf{D}\widehat{\mathbf{B}}\mathbf{F}$ *is unique if and only if* $\mathcal{C}(\mathbf{D}') \subseteq \mathcal{C}(\mathbf{A}')$ *and* $\mathcal{C}(\mathbf{F}) \subseteq \mathcal{C}(\mathbf{C})$.

Example 2.3 *(Example 2.2 continued) The* $\mathbf{A}$ *matrix in Example 2.1 has full column rank, i.e.* $\rho(\mathbf{A}) = 4$*. One way to verify this is to calculate the determinant of the sub-matrix obtained by omitting the last* $p-4$ *rows, i.e. if this matrix is denoted* $\mathbf{A}_{sub}$,

$$\mathbf{A}_{sub} = \begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ 1 & t_2 & t_2^2 & t_2^3 \\ 1 & t_3 & t_3^2 & t_3^3 \\ 1 & t_4 & t_4^2 & t_4^3 \end{pmatrix}.$$

Then $det(\mathbf{A}_{sub})$ *is the well-known so called Vandermonde determinant, which means that* $det(\mathbf{A}_{sub}) = \prod_{i<j}(t_j - t_i)$, $i = 1,2,3,\ j = 2,3,4$*. Since* $t_j - t_i > 0$ *for* $i < j$, $det(\mathbf{A}_{sub}) > 0$*, so that* $\mathbf{A}_{sub}$ *is non-singular with* $\rho(\mathbf{A}_{sub}) = 4$*, which also implies* $\rho(\mathbf{A}) = 4$*. More generally, if the expected response is a polynomial function of degree* $q-1$, $q \leq p$*, then the rank of* $\mathbf{A}$ *equals* q.

The ML estimators of the parameters in GCM2 for arbitrary m can be found in von Rosen (1989). As mentioned GCM2 becomes useful when linear restrictions on the parameters exist in GCM1. To see why, suppose $\mathbf{DBF} = \mathbf{0}$

for known matrices $\mathbf{D}$ and $\mathbf{F}$. This is a homogeneous system of equations in $\mathbf{B}$ which has the general solution $\mathbf{B} = (\mathbf{D}')^o\boldsymbol{\Theta}_1 + \mathbf{D}'\boldsymbol{\Theta}_2\mathbf{F}^{o'}$, where $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ are new parameter matrices. If the expression for $\mathbf{B}$ is inserted in GCM1 we obtain

$$\mathbf{X} = \mathbf{A}(\mathbf{D}')^o\boldsymbol{\Theta}_1\mathbf{C} + \mathbf{A}\mathbf{D}'\boldsymbol{\Theta}_2\mathbf{F}^{o'}\mathbf{C} + \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{E}.$$

However, since $\mathcal{C}(\mathbf{C}'\mathbf{F}^o) \subseteq \mathcal{C}(\mathbf{C}')$ this model is in fact a GCM2 with $m = 2$ and unknown parameters $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$. Thus, the ML estimators of the parameters of $\mathbf{B}$ and $\boldsymbol{\Sigma}$ can be derived. The results are presented in the next lemma.

Lemma 2.2 *Suppose* $\mathbf{DBF} = \mathbf{0}$ *in the GCM1, where* $\mathbf{D}$*:* $s \times q$*,* $\mathbf{F}$*:* $k \times t$ *are known matrices. Then*

$$\begin{aligned}\widehat{\mathbf{B}} &= (\mathbf{D}')^o\widehat{\boldsymbol{\Theta}}_1 + \mathbf{D}'\widehat{\boldsymbol{\Theta}}_2\mathbf{F}^{o'},\\ n\widehat{\boldsymbol{\Sigma}} &= (\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C})(\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C})',\end{aligned}$$

where

$$\begin{aligned}\widehat{\boldsymbol{\Theta}}_2 &= (\mathbf{D}\mathbf{A}'\mathbf{T}_1'\mathbf{S}_2^{-1}\mathbf{T}_1\mathbf{A}\mathbf{D}')^-\mathbf{D}\mathbf{A}'\mathbf{T}_1'\mathbf{S}_2^{-1}\mathbf{T}_1\mathbf{X}\mathbf{C}'\mathbf{F}^o(\mathbf{F}^{o'}\mathbf{C}\mathbf{C}'\mathbf{F}^o)^-\\ &\quad +(\mathbf{D}\mathbf{A}'\mathbf{T}_1')^o\mathbf{Z}_{11} + \mathbf{D}\mathbf{A}'\mathbf{T}_1'\mathbf{Z}_{12}(\mathbf{F}^{o'}\mathbf{C})^{o'},\\ \widehat{\boldsymbol{\Theta}}_1 &= ((\mathbf{D}')^{o'}\mathbf{A}'\mathbf{S}_1^{-1}\mathbf{A}(\mathbf{D}')^o)^-(\mathbf{D}')^{o'}\mathbf{A}'\mathbf{S}_1^{-1}(\mathbf{X} - \mathbf{A}\mathbf{D}'\widehat{\boldsymbol{\Theta}}_2\mathbf{F}^{o'}\mathbf{C})\mathbf{C}'(\mathbf{C}\mathbf{C}')^-\\ &\quad +((\mathbf{D}')^{o'}\mathbf{A}')^o\mathbf{Z}_{21} + (\mathbf{D}')^{o'}\mathbf{A}'\mathbf{Z}_{22}\mathbf{C}^{o'},\\ \mathbf{T}_1 &= \mathbf{I} - \mathbf{A}(\mathbf{D}')^o((\mathbf{D}')^{o'}\mathbf{A}'\mathbf{S}_1^{-1}\mathbf{A}(\mathbf{D}')^o)^-(\mathbf{D}')^{o'}\mathbf{A}'\mathbf{S}_1^{-1},\\ \mathbf{S}_1 &= \mathbf{X}(\mathbf{I} - \mathbf{C}'(\mathbf{C}\mathbf{C}')^-\mathbf{C})\mathbf{X}',\\ \mathbf{S}_2 &= \mathbf{S}_1 + \mathbf{T}_1\mathbf{X}\mathbf{C}'(\mathbf{C}\mathbf{C}')^-\mathbf{C}(\mathbf{I} - \mathbf{C}'\mathbf{F}^o(\mathbf{F}^{o'}\mathbf{C}\mathbf{C}'\mathbf{F}^o)^-\mathbf{F}^{o'}\mathbf{C})\mathbf{C}'(\mathbf{C}\mathbf{C}')^-\mathbf{C}\mathbf{X}'\mathbf{T}_1',\end{aligned}$$

$\mathbf{S}_1$ *is assumed to be p.d. and* $\mathbf{Z}_{ij}$*,* $i, j = 1, 2$*, are arbitrary matrices.*

In the next lemma uniqueness conditions are given for $\widehat{\mathbf{B}}$ and linear combinations $\mathbf{K}\widehat{\mathbf{B}}\mathbf{L}$ as well as the dispersion matrix of $\widehat{\mathbf{B}}$. For proofs see von Rosen (1990, 1995).

Lemma 2.3 *Suppose that the same conditions as in Lemma 2.2 hold.*

(i) *The estimator* $\widehat{\mathbf{B}}$ *is unique if and only if*

$$\begin{aligned}&\rho(\mathbf{C}) = k,\\ &\mathcal{C}(\mathbf{D}') \cap \mathcal{C}(\mathbf{A}')^\perp = \{\mathbf{0}\},\\ &\mathcal{C}(\mathbf{D}') \cap \{\mathcal{C}(\mathbf{D}')^\perp + \mathcal{C}(\mathbf{A}')^\perp\} = \{\mathbf{0}\}.\end{aligned}$$

(ii) *The linear combinations* $\mathbf{K}\widehat{\mathbf{B}}\mathbf{L}$, *where* $\mathbf{K}$ *and* $\mathbf{L}$ *are known matrices, are unique if and only if*

$$\begin{aligned}
\mathcal{C}(\mathbf{L}) &\subseteq \mathcal{C}(\mathbf{C}),\\
\mathcal{C}((\mathbf{D}')^{o'}\mathbf{K}') &\subseteq \mathcal{C}((\mathbf{D}')^{o'}\mathbf{A}'),\\
\mathcal{C}(\mathbf{F}^{o'}\mathbf{L}) &\subseteq \mathcal{C}(\mathbf{F}^{o'}\mathbf{C}),\\
\mathcal{C}(\mathbf{D}\mathbf{P}_1'\mathbf{K}') &\subseteq \mathcal{C}(\mathbf{D}\mathbf{P}_1'\mathbf{A}'),
\end{aligned}$$

where

$$\mathbf{P}_1 = \mathbf{I} - (\mathbf{D}')^o((\mathbf{D}')^{o'}\mathbf{A}'\mathbf{A}(\mathbf{D}')^o)^-(\mathbf{D}')^{o'}\mathbf{A}'\mathbf{A}.$$

(iii) *Let*

$$\begin{aligned}
c_1 &= \frac{n-k-1}{n-k-p+q-1},\\
c_2 &= \tfrac{q(n-k+\rho(\mathbf{F})-1)}{(n-k-p+q-1)(n-k+\rho(\mathbf{F})-p+2q-1)},\\
c_3 &= \tfrac{n-k+\rho(\mathbf{F})-1}{n-k+\rho(\mathbf{F})-p+q-1}.
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{L}_1 &= (\mathbf{D}')^o((\mathbf{D}')^{o'}\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A}(\mathbf{D}')^o)^-(\mathbf{D}')^{o'},\\
\mathbf{L}_2 &= (\mathbf{I} - (\mathbf{D}')^o((\mathbf{D}')^{o'}\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A}(\mathbf{D}')^o)^-(\mathbf{D}')^{o'}\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A})\\
&\quad\times\mathbf{D}'(\mathbf{D}\mathbf{A}'(\mathbf{A}(\mathbf{D}')^o)^o((\mathbf{A}(\mathbf{D}')^o)^{o'}\mathbf{\Sigma}(\mathbf{A}(\mathbf{D}')^o)^o)^{-1}(\mathbf{A}(\mathbf{D}')^o)^{o'}\mathbf{A}\mathbf{D})^-\mathbf{D}\\
&\qquad\times(\mathbf{I} - \mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A}(\mathbf{D}')^o((\mathbf{D}')^{o'}\mathbf{A}'\mathbf{\Sigma}^{-1}\mathbf{A}(\mathbf{D}')^o)^-(\mathbf{D}')^{o'}).
\end{aligned}$$

Then,

$$D[\widehat{\mathbf{B}}] = \mathbf{R}_1 + \mathbf{R}_2,$$

where

$$\begin{aligned}
\mathbf{R}_1 &= c_1(\mathbf{C}\mathbf{C}')^{-1} \otimes \mathbf{L}_1,\\
\mathbf{R}_2 &= \mathbf{F}^o(\mathbf{F}^{o'}\mathbf{C}\mathbf{C}'\mathbf{F}^o)^-\mathbf{F}^{o'} \otimes (3c_2\mathbf{L}_1 + (c_3+2)\mathbf{L}_2).
\end{aligned}$$

As in Lemma 2.1 (iii) we could have presented unbiased estimators of $D[\widehat{\mathbf{B}}]$ but they will be rather lengthy. However, $\widehat{\mathbf{\Sigma}}$ is a consistent estimator of $\mathbf{\Sigma}$ and therefore in order to find reasonable estimators of $D[\widehat{\mathbf{B}}]$ we replace $\mathbf{\Sigma}$ by $\widehat{\mathbf{\Sigma}}$. Observe that the correction performed in Lemma 2.1 (iii), i.e. replacing c_1 by c_{11} does not have any strong implication.

2.3 Cross-over designs and mixed linear models

For reviews of the cross-over design see Jones and Kenward (1987) and Senn (1993). A common practice is that repeated measurements are collected for a response variable in each period from each subject. Some examples are

1. Systolic blood pressure measured 2, 4, 8, 24 and 48 hours after administration of a drug.
2. Lamb weights taken at eight two-weekly intervals during an experiment on growth rate.
3. Concentration of a test drug in the blood measured every ten minutes after administration of the drug.

The term repeated measures cross-over design is used for this type of design. We will assume that measurements are made at equivalent times for each subject and for each period, although this is not always necessary. Several different methods to analyze this design have been proposed. If a uniform covariance structure (compound symmetry) between repeated measurements is assumed, ANOVA-methods can be used. However, since the repeated measurements within periods are made during a short time interval, the uniform structure assumption often does not make sense. Another disadvantage with ANOVA is that the change of the response variable across the period of measurement is usually modelled with a general time effect and time by treatment interaction, whereas a model of the response by a time-dependent function, e.g. a linear or logarithmic growth rate in Example 1 above, would be preferred. Putt and Chinchilli (1999) present a mixed effects model that can model time-dependent changes, which is based on the random coefficients growth curve model by Rao (1965). Since this model probably is the most commonly applied approach today, we will summarize it briefly. The model, which in principle can be applied to all types of cross-over designs, is written

$$\mathbf{Y}_{ijkl} = \mathbf{w}_{ijkl}(\boldsymbol{\beta}_{ikl} + \boldsymbol{\delta}_{ijk}) + \boldsymbol{\epsilon}_{ijkl},$$

where i, j, k and l indexes sequence, subject in sequence i, treatment and replicate of treatment k within sequence i, respectively. The response vector, $\mathbf{Y}_{ijkl}$ is a q-vector of measurements of the jth subject on the ith sequence for the lth replicate of treatment k, $\mathbf{w}_{ijkl}$ is a q×r design matrix, $\boldsymbol{\beta}_{ikl}$ and $\boldsymbol{\delta}_{ijk}$ are the r-vectors of fixed and random effects, respectively, and $\boldsymbol{\epsilon}_{ijkl}$ is a q-vector of random errors. The fixed effect $\boldsymbol{\beta}_{ikl}$ is composed of parameters representing mean and nuisance effects. If the random subject effect $\boldsymbol{\delta}_{ijk}$

is written $\boldsymbol{\delta}_{ij} = (\boldsymbol{\delta}'_{ij1}, \ldots, \boldsymbol{\delta}'_{ijt})'$, where t is the number of treatments, the between-subject dispersion matrix is the tr×tr-matrix

$$D[\boldsymbol{\delta}_{ij}] = \begin{pmatrix} \boldsymbol{\Omega}_{11} & \ldots & \boldsymbol{\Omega}_{1t} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Omega}_{1t} & \ldots & \boldsymbol{\Omega}_{tt} \end{pmatrix},$$

where $\boldsymbol{\Omega}_{kk'} = (\omega_{kk'}^{(mm')})$ and $\omega_{kk'}^{(mm')}$ is the between-subject covariance for the mth and m'th location parameter on treatment k and k', respectively. The within-subject dispersion matrix is

$$D[\boldsymbol{\epsilon}_{ijkl}] = \begin{pmatrix} \sigma_k^{(11)} & \ldots & \sigma_k^{(1q)} \\ \vdots & \ddots & \vdots \\ \sigma_k^{(q1)} & \ldots & \sigma_k^{qq} \end{pmatrix}.$$

In order to obtain estimators of the fixed effects, the generalized least square (GLS) method is used. In the case of normally distributed data, the GLS estimator of the fixed effects will be equal to the maximum likelihood ML estimator. However, the within-subject dispersion matrix must be known in order to calculate the GLS estimate, except in the case of independent random error, i.e. all $\sigma_k^{(mm')} = 0$, $m \neq m'$, $k = 1, \ldots, t$. When this is not the case, the estimated generalized least square (EGLS) method may be used (see Putt and Chinchilli, 1999) . For independent random errors the GLS estimator will be equal to the ordinary least square (OLS) estimator. In Putt and Chinchilli (1999), all models that were evaluated assumed independent random errors and an unstructured matrix for the random subject effects. Explicit (closed form) expressions for the ML estimates of the variance components were obtained for designs which are uniform within sequence (i.e. for each subject, each treatment appears in the same number of periods), there are no covariates and where every subject has the same design matrix.

3 Growth curve models for the analysis of cross-over designs

3.1 Introduction

In Section 3.2 – 3.4, a two-sequence two-period repeated measures cross-over design is considered. Firstly, the model is defined in terms of a multivariate linear normal model and the maximum likelihood estimators of the parameters are derived. Secondly, a transformation of the problem is considered, that under certain additional assumptions of the dispersion matrix, simplifies the derivation of the maximum likelihood estimators of the parameters. In Section 3.5, the results in Section 3.2 – 3.4 are extended to a two-sequence four-period design. Lastly, the models are illustrated numerically by calculating the ML estimates for a cross-over study that compares two different insulin mixtures in rabbits.

3.2 Description of the model for a two-sequence two-period design

We will assume that a time series of p ($p \geq 1$) measurements of a response variable are collected within periods for each subject, and that the measurements are made at equivalent times for each subject in each period. The $2p$-vectors of measurements from subjects are assumed to be independent observations from a multivariate normal distribution. We will also assume that the variance is equal for the two treatments. The expected response variable could be a polynomial function of time with unknown coefficients, or some other unknown linear combination of a time-dependent function. The aim of a study could be to compare the time course of changes in the response variable for treatments. It will be shown that the model can be considered as a GCM1 and consequently results given in Section 2.2 can be used.

Let n_1 and n_2 be the number of subjects in treatment sequence 1 and 2, respectively, and $n = n_1 + n_2$ the total number of subjects. Let $t_1, \ldots, t_p$ be the time points of measurements, which are assumed to be the same for each subject. Let $(\mathbf{X}'_{i1l} : \mathbf{X}'_{i2l})'$ be the random observable $2p \times 1$ vector for subject i, $i = 1, \ldots, n_1 + n_2$, where the vector has been partitioned by period and index l, $l = 1, 2$ represents treatment sequence. Thus, $l = 1$ for $i = 1, \ldots, n_1$ and $l = 2$ for $i = n_1 + 1, \ldots, n_1 + n_2$. We assume that $(\mathbf{X}'_{i1l} : \mathbf{X}'_{i2l})'$ will follow the model

$$\begin{pmatrix} \mathbf{X}_{i1l} \\ \mathbf{X}_{i2l} \end{pmatrix} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix} \begin{pmatrix} \mathbf{B}_{1l} \\ \mathbf{B}_{2l} \end{pmatrix} + \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}^{\frac{1}{2}} \mathbf{e}_i,$$

where the $\mathbf{e}_i$:s are independently distributed according to $N_{2p}(\mathbf{0}, \mathbf{I})$, $\mathbf{A}_1$: $p \times q$ is the within subject known design matrix with $\rho(\mathbf{A}_1) = q$ and $q \leq p$; $\mathbf{B}_{jl}$ is the q-vector of unknown parameters for period $j = 1, 2$ and treatment sequence $l = 1, 2$; $\mathbf{\Sigma}_{jj}$ is the dispersion matrix for period j; and $\mathbf{\Sigma}_{12}$ is the matrix of covariances between components in period 1 and 2. This means that

$(\mathbf{X}'_{i1l} : \mathbf{X}'_{i2l})'$ are independently multivariate normally distributed with mean

$$\begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix} \begin{pmatrix} \mathbf{B}_{1l} \\ \mathbf{B}_{2l} \end{pmatrix},$$

and dispersion matrix

$$\begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}.$$

In a standard set-up (see Jones and Kenward, 1987) $\mathbf{B}_{jl}$ comprises a general mean effect, treatment effect, period effect and a simple carry-over effect. Thus,

$$\begin{aligned} \mathbf{B}_{11} &= \boldsymbol{\mu} + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_1, \\ \mathbf{B}_{12} &= \boldsymbol{\mu} + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_1, \\ \mathbf{B}_{21} &= \boldsymbol{\mu} + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_2 + \boldsymbol{\lambda}_1, \\ \mathbf{B}_{22} &= \boldsymbol{\mu} + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_2 + \boldsymbol{\lambda}_2, \end{aligned}$$

where the terms in $\mathbf{B}_{jl}$ are:

$$\begin{aligned} \boldsymbol{\mu} &= (\mu_1, \ldots, \mu_q)' : && \text{a general mean effect,} \\ \boldsymbol{\tau}_r &= (\tau_{1r}, \ldots, \tau_{qr})' : && \text{the effect of treatment } r,\ r = 1, 2, \\ \boldsymbol{\pi}_j &= (\pi_{1j}, \ldots, \pi_{qj})' : && \text{the effect of period } j,\ j = 1, 2, \\ \boldsymbol{\lambda}_r &= (\lambda_{1r}, \ldots, \lambda_{qr})' : && \text{the carry-over effect from treatment } r \end{aligned}$$

However, it should be noted that the inclusion of a simple carry-over effect in a AB:BA cross-over design is controversial (see Senn, 1993). Therefore, later on a model where $\mathbf{B}_{jl}$ comprises only of a general mean-, treatment- and period effect will also be evaluated. Let the observation matrix $\mathbf{X}$ be the $2p \times n$ matrix with $(\mathbf{X}'_{i1l} : \mathbf{X}'_{i2l})'$ as columns. Then, the model in matrix notation becomes

$$\mathbf{X} = \mathbf{ABC} + \mathbf{\Sigma}^{1/2}\mathbf{E}, \tag{1}$$

where

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{1}'_{n_1} & \mathbf{0}'_{n_2} \\ \mathbf{0}'_{n_1} & \mathbf{1}'_{n_2} \end{pmatrix}, \\ \mathbf{B} &= \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu} + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_1 & \boldsymbol{\mu} + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_1 \\ \boldsymbol{\mu} + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_2 + \boldsymbol{\lambda}_1 & \boldsymbol{\mu} + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_2 + \boldsymbol{\lambda}_2 \end{pmatrix}, \\ \mathbf{\Sigma} &= \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{pmatrix}, \end{aligned}$$

and the elements e_{ij} of $\mathbf{E} : 2p \times n$ are iid $N(0, 1)$. The model is now given in the same form as the GCM1. However, there is a difference in the definition of the parameter matrix $\mathbf{B}$ in our case on the one hand, and in the GCM1 on the other hand. The main interest in our case lies in the estimators of $\boldsymbol{\mu}$, $\boldsymbol{\tau}_r$, $\boldsymbol{\pi}_j$ and $\boldsymbol{\lambda}_r$, and in linear combinations of these. This problem is solved in the next section.

3.3 Maximum likelihood estimators of the parameters

Maximum likelihood estimators of the parameters are presented in the next theorem.

Theorem 3.1 *Suppose that we have a two-sequence two-period repeated measures AB:BA cross-over design with n_1 and n_2 subjects in treatment sequence 1 and 2, respectively. Suppose that a response variable is measured at time points $t_1, \ldots, t_p$ within periods for each subject, and that the 2p-vectors of outcomes from subjects are independent observations from a multivariate normal distribution with a positive definite dispersion matrix, $\boldsymbol{\Sigma}$. Suppose $n = n_1 + n_2 \geq 2p + 2$. Let further $\mathbf{A}$, $\mathbf{A}_1$, $\mathbf{B}$ and $\mathbf{C}$ be defined as in Section 3.2.*

(i) *If the constraints*

$$\begin{aligned} \boldsymbol{\tau}_1 + \boldsymbol{\tau}_2 &= \mathbf{0}, & (2)\\ \boldsymbol{\pi}_1 + \boldsymbol{\pi}_2 &= \mathbf{0}, & (3)\\ \boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 &= \mathbf{0} & (4) \end{aligned}$$

are imposed on the parameters, and, for convenience, the parameters are renamed as $\boldsymbol{\tau} = \boldsymbol{\tau}_1$, $\boldsymbol{\pi} = \boldsymbol{\pi}_1$ and $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1$, then the maximum likelihood estimator of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$ and $\boldsymbol{\lambda}$ equal

$$\widehat{\boldsymbol{\mu}} = \frac{1}{4}\left(\mathbf{I} : \mathbf{I}\right)\widehat{\mathbf{B}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{4}(\widehat{\mathbf{B}}_{11} + \widehat{\mathbf{B}}_{21} + \widehat{\mathbf{B}}_{12} + \widehat{\mathbf{B}}_{22}), \tag{5}$$

$$\widehat{\boldsymbol{\tau}} = \frac{1}{2}\left(\mathbf{I} : \mathbf{0}\right)\widehat{\mathbf{B}}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{2}(\widehat{\mathbf{B}}_{11} - \widehat{\mathbf{B}}_{12}), \tag{6}$$

$$\widehat{\boldsymbol{\pi}} = \frac{1}{4}\left(\mathbf{I} : -\mathbf{I}\right)\widehat{\mathbf{B}}\begin{pmatrix}1\\1\end{pmatrix} = \frac{1}{4}(\widehat{\mathbf{B}}_{11} - \widehat{\mathbf{B}}_{21} + \widehat{\mathbf{B}}_{12} - \widehat{\mathbf{B}}_{22}), \tag{7}$$

$$\widehat{\boldsymbol{\lambda}} = \frac{1}{2}\left(\mathbf{I} : \mathbf{I}\right)\widehat{\mathbf{B}}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{1}{2}(\widehat{\mathbf{B}}_{11} + \widehat{\mathbf{B}}_{21} - \widehat{\mathbf{B}}_{12} - \widehat{\mathbf{B}}_{22}), \tag{8}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are defined in Section 3.1, and $\widehat{\mathbf{B}}$ and $\widehat{\boldsymbol{\Sigma}}$ are given by Lemma 2.1.

(ii) *The maximum likelihood estimators of* $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$ *and* $\boldsymbol{\lambda}$ *are unique if and only if* $\rho(\mathbf{A}_1) = q$.

Proof (i) First note that the model satisfies the conditions of the GCM1 (see Definition 2.1) and therefore the ML estimators of $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are given by Lemma 2.1. If the constraints in (2)-(4) are used, the parameter matrix $\mathbf{B}$ equals

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}+\boldsymbol{\tau}+\boldsymbol{\pi} & \boldsymbol{\mu}-\boldsymbol{\tau}+\boldsymbol{\pi} \\ \boldsymbol{\mu}-\boldsymbol{\tau}-\boldsymbol{\pi}+\boldsymbol{\lambda} & \boldsymbol{\mu}+\boldsymbol{\tau}-\boldsymbol{\pi}-\boldsymbol{\lambda} \end{pmatrix}. \tag{9}$$

The relation in (9) represents a linear transformation, L: $\mathbb{R}^{4q} \to \mathbb{R}^{4q}$, of $\mathbf{v} = (\boldsymbol{\mu}' : \boldsymbol{\tau}' : \boldsymbol{\pi}' : \boldsymbol{\lambda}')'$ to $\mathbf{w} = \text{vec}(\mathbf{B}) = (\mathbf{B}_{11}' : \mathbf{B}_{21}' : \mathbf{B}_{12}' : \mathbf{B}_{22}')'$ which can be written as $L(\mathbf{v}) = \mathbf{M}\mathbf{v}$, where

$$\mathbf{M} = \begin{pmatrix} \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} & -\mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} & \mathbf{I} & \mathbf{0} \\ \mathbf{I} & \mathbf{I} & -\mathbf{I} & -\mathbf{I} \end{pmatrix} : \ 4q \times 4q.$$

Since the inverse of $\mathbf{M}$ exists, L is an one-to-one linear transformation. Therefore, $L^{-1}(\widehat{\mathbf{w}}) = \widehat{L^{-1}(\mathbf{w})} = \widehat{\mathbf{M}^{-1}\mathbf{w}} = \widehat{\mathbf{v}}$, where $\widehat{\ }$ denotes the ML estimator. The inverse of $\mathbf{M}$ equals

$$\mathbf{M}^{-1} = \frac{1}{4} \begin{pmatrix} \mathbf{I} & \mathbf{I} & \mathbf{I} & \mathbf{I} \\ 2\mathbf{I} & \mathbf{0} & -2\mathbf{I} & \mathbf{0} \\ \mathbf{I} & -\mathbf{I} & \mathbf{I} & -\mathbf{I} \\ 2\mathbf{I} & 2\mathbf{I} & -2\mathbf{I} & -2\mathbf{I} \end{pmatrix},$$

which gives the ML estimators of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$ and $\boldsymbol{\lambda}$.

(ii) Since $\rho(\mathbf{A}_1) = q$ if and only if $\rho(\mathbf{A}) = 2q$, the results follow from Lemma 2.1(ii). Observe that $\rho(\mathbf{C}) = 2$ follows from the definition of the design. □

In the next theorem, a model is considered where no carry-over effects are supposed to exist from treatments, i.e. when $\boldsymbol{\lambda}_1 = \boldsymbol{\lambda}_2 = \boldsymbol{\lambda} = \mathbf{0}$. The ML estimators of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$ will have the same expressions as in Theorem 3.1, but $\widehat{\mathbf{B}}$ will now be given by Lemma 2.2.

Theorem 3.2 *Suppose that the same conditions as in Theorem 3.1 hold. In addition, suppose* $\boldsymbol{\lambda} = \mathbf{0}$. *Then, the maximum likelihood estimators of* $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$ *are given by (5)-(7), where* $\widehat{\mathbf{B}}$ *and* $\widehat{\boldsymbol{\Sigma}}$ *are presented in Lemma 2.2.*

Proof From the ML estimator of $\boldsymbol{\lambda}$ in Theorem 3.1, we see that if $\mathbf{D}$: $q \times 2q$ and $\mathbf{F}$: 2×1 are defined by

$$\mathbf{D} = \begin{pmatrix} \mathbf{I}_q : & \mathbf{I}_q \end{pmatrix} \quad \text{and} \quad \mathbf{F} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Then $\mathbf{DBF} = 2\boldsymbol{\lambda}$ and thus $\boldsymbol{\lambda} = \mathbf{0}$ if and only if $\mathbf{DBF} = \mathbf{0}$. The model now satisfies the conditions of GCM1 with linear restrictions $\mathbf{DBF} = \mathbf{0}$. Thus, the ML estimators of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$ and $\boldsymbol{\pi}$ will have the same expression as in Theorem 3.1, except that $\widehat{\mathbf{B}}$ and $\widehat{\boldsymbol{\Sigma}}$ now are based on Lemma 2.2. $\square$

The uniqueness of $\widehat{\mathbf{B}}$ and $\widehat{\boldsymbol{\mu}}$, $\widehat{\boldsymbol{\tau}}$, $\widehat{\boldsymbol{\pi}}$ depends on the structure of $\mathbf{A}$ and conditions are given by Lemma 2.3. In particular, if $\rho(\mathbf{A}) = 2q$ then $\widehat{\boldsymbol{\tau}}$ will be unique, see Åsenblad (2001). Note that a "natural estimator" of $\boldsymbol{\tau}$ is $\frac{1}{4}((\widehat{\mathbf{B}}_{11} - \widehat{\mathbf{B}}_{12}) - (\widehat{\mathbf{B}}_{21} - \widehat{\mathbf{B}}_{22}))$. This estimator could be seen as a generalization of the standard period-adjusted estimator of $\boldsymbol{\tau}$ in the standard AB:BA design. However, due to the linear restrictions of $\mathbf{B}$ imposed by the assumption $\boldsymbol{\lambda} = \mathbf{0}$, the "natural estimator" of $\boldsymbol{\tau}$ will be the same as the estimator given by Theorem 3.2.

3.4 A model with a structure on the dispersion matrix

A model with structure on the dispersion matrix is considered that, together with a linear transformation of the model, will simplify the problem of deriving maximum likelihood estimators. Let $\boldsymbol{\Gamma}$ be the $2p \times 2p$ non-singular matrix

$$\boldsymbol{\Gamma} = \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{pmatrix}. \tag{10}$$

Then the first p rows of $\boldsymbol{\Gamma}\mathbf{X}$ will consist of the sum of the two p-vectors in period 1 and 2, and the last p rows of the difference between the same p-vectors. That is,

$$\boldsymbol{\Gamma}\mathbf{X} = \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{X}_{21} & \mathbf{X}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{X}_{11} + \mathbf{X}_{21} & \mathbf{X}_{12} + \mathbf{X}_{22} \\ \mathbf{X}_{11} - \mathbf{X}_{21} & \mathbf{X}_{12} - \mathbf{X}_{22} \end{pmatrix},$$

where $\mathbf{X}_{ij}$: $p \times n_j$ is the submatrix of $\mathbf{X}$ representing period i and treatment sequence j. The expectation of $\boldsymbol{\Gamma}\mathbf{X}$ is given by

$$\begin{aligned} E[\boldsymbol{\Gamma}\mathbf{X}] = \boldsymbol{\Gamma}E[\mathbf{X}] = \boldsymbol{\Gamma}\mathbf{ABC} &= \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix} \mathbf{BC} \\ &= \mathbf{A} \begin{pmatrix} \mathbf{B}_{11} + \mathbf{B}_{21} & \mathbf{B}_{12} + \mathbf{B}_{22} \\ \mathbf{B}_{11} - \mathbf{B}_{21} & \mathbf{B}_{12} - \mathbf{B}_{22} \end{pmatrix} \mathbf{C}. \end{aligned}$$

Moreover, the dispersion matrix for the columns of $\mathbf{\Gamma X}$ equals

$$
\begin{aligned}
D[(\mathbf{\Gamma X})_i] &= \mathbf{\Gamma} D[\mathbf{X}_i]\mathbf{\Gamma}' = \mathbf{\Gamma\Sigma\Gamma}' \\
&= \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{\Sigma}_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}' & \mathbf{\Sigma}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{I} & -\mathbf{I} \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{\Sigma}_{11} + \mathbf{\Sigma}_{12}' + \mathbf{\Sigma}_{12} + \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{11} + \mathbf{\Sigma}_{12}' - \mathbf{\Sigma}_{12} - \mathbf{\Sigma}_{22} \\ \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}' + \mathbf{\Sigma}_{12} - \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{11} - \mathbf{\Sigma}_{12}' - \mathbf{\Sigma}_{12} + \mathbf{\Sigma}_{22} \end{pmatrix},
\end{aligned}
$$

where subscript i represents column i of $\mathbf{X}$.

If we make the additional assumptions that the dispersion matrix is equal for period 1 and 2 and that the covariances between period 1 and 2 are equal, that is, $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22}$ and $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{12}'$, respectively, then the dispersion matrix for the columns of $\mathbf{\Gamma X}$ becomes

$$
D[(\mathbf{\Gamma X})_i] = \begin{pmatrix} 2\mathbf{\Sigma}_{11} + 2\mathbf{\Sigma}_{12} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{\Sigma}_{11} - 2\mathbf{\Sigma}_{12} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{pmatrix},
$$

which implies that the column vectors of $\mathbf{X}_{1j} + \mathbf{X}_{2j}$ and $\mathbf{X}_{1j} - \mathbf{X}_{2j}$ are independently multivariate normally distributed, $j = 1, 2$. This result is important and will be used to obtain estimators of the parameters in a convenient way.

The ML estimators of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$, $\boldsymbol{\lambda}$ and $\mathbf{\Sigma}$ are presented in the next theorem. In the following, it will be convenient to use the notation: $\mathbf{Y}_{11} = \mathbf{X}_{11} + \mathbf{X}_{21}$, $\mathbf{Y}_{12} = \mathbf{X}_{12} + \mathbf{X}_{22}$, $\mathbf{Y}_{21} = \mathbf{X}_{11} - \mathbf{X}_{21}$, $\mathbf{Y}_{22} = \mathbf{X}_{12} - \mathbf{X}_{22}$, $\mathbf{H}_{11} = \mathbf{B}_{11} + \mathbf{B}_{21}$, $\mathbf{H}_{12} = \mathbf{B}_{12} + \mathbf{B}_{22}$, $\mathbf{H}_{21} = \mathbf{B}_{11} - \mathbf{B}_{21}$ and $\mathbf{H}_{22} = \mathbf{B}_{12} - \mathbf{B}_{22}$. Furthermore, the notation $()'$ will be used to denote multiplication with the preceding factor transposed.

Theorem 3.3 *Suppose that the same conditions as in Theorem 3.1 hold. In addition, suppose that $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22}$ and $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{12}'$. Then*

(i) *The maximum likelihood estimators of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$, $\boldsymbol{\lambda}$, $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$, $\mathbf{\Sigma}_{11}$ and $\mathbf{\Sigma}_{12}$ equal*

$$
\begin{aligned}
\widehat{\boldsymbol{\mu}} &= \frac{1}{4}(\widehat{\mathbf{H}}_{11} + \widehat{\mathbf{H}}_{12}), \\
\widehat{\boldsymbol{\lambda}} &= \frac{1}{2}(\widehat{\mathbf{H}}_{11} - \widehat{\mathbf{H}}_{12}), \\
\widehat{\boldsymbol{\pi}} &= \frac{1}{4}(\widehat{\mathbf{H}}_{21} + \widehat{\mathbf{H}}_{22}), \\
\widehat{\boldsymbol{\tau}} &= \frac{1}{4}(\widehat{\mathbf{H}}_{11} - \widehat{\mathbf{H}}_{12} + \widehat{\mathbf{H}}_{21} - \widehat{\mathbf{H}}_{22}),
\end{aligned}
$$

$$
\begin{aligned}
n\widehat{\mathbf{\Sigma}}_1 &= ((\mathbf{Y}_{11} : \mathbf{Y}_{12}) - \mathbf{A}_1(\widehat{\mathbf{H}}_{11} : \widehat{\mathbf{H}}_{12})\mathbf{C})()', \\
n\widehat{\mathbf{\Sigma}}_2 &= ((\mathbf{Y}_{21} : \mathbf{Y}_{22}) - \mathbf{A}_1(\widehat{\mathbf{H}}_{21} : \widehat{\mathbf{H}}_{22})\mathbf{C})()', \\
\widehat{\mathbf{\Sigma}}_{11} &= \frac{1}{4}(\widehat{\mathbf{\Sigma}}_1 + \widehat{\mathbf{\Sigma}}_2), \\
\widehat{\mathbf{\Sigma}}_{12} &= \frac{1}{4}(\widehat{\mathbf{\Sigma}}_1 - \widehat{\mathbf{\Sigma}}_2),
\end{aligned}
$$

where $(\widehat{\mathbf{H}}_{11} : \widehat{\mathbf{H}}_{12})$ *and* $(\widehat{\mathbf{H}}_{21} : \widehat{\mathbf{H}}_{22})$ *are obtainable from a GCM1 with mean* $\mathbf{A}_1(\mathbf{H}_{11} : \mathbf{H}_{12})\mathbf{C}$ *and* $\mathbf{A}_1(\mathbf{H}_{21} : \mathbf{H}_{22})\mathbf{C}$*, respectively.*

(ii) *The maximum likelihood estimators are unique if and only if* $\rho(\mathbf{A}_1) = q$*.*

Proof (i) The model (1) becomes, after transformation by $\mathbf{\Gamma}$ given in (10),

$$
\begin{pmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} \end{pmatrix} = \mathbf{A} \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \mathbf{C} + \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{pmatrix}^{\frac{1}{2}} \mathbf{E}.
$$

Since the column vectors of $\mathbf{X}_{ij} + \mathbf{X}_{ij}$ and $\mathbf{X}_{ij} - \mathbf{X}_{ij}$ are independently multivariate normally distributed, the model can be divided into two models

$$
\begin{aligned}
(\mathbf{Y}_{11} : \mathbf{Y}_{12}) &= \mathbf{A}_1(\mathbf{H}_{11} : \mathbf{H}_{12})\mathbf{C} + \mathbf{\Sigma}_1^{\frac{1}{2}}\mathbf{E}_1, & (11) \\
(\mathbf{Y}_{21} : \mathbf{Y}_{22}) &= \mathbf{A}_1(\mathbf{H}_{21} : \mathbf{H}_{22})\mathbf{C} + \mathbf{\Sigma}_2^{\frac{1}{2}}\mathbf{E}_2. & (12)
\end{aligned}
$$

Thus, the models in (11) and (12) are GCM1 with mean $\mathbf{A}_1(\mathbf{H}_{11} : \mathbf{H}_{12})\mathbf{C})$ and $\mathbf{A}_1(\mathbf{H}_{21} : \mathbf{H}_{22})\mathbf{C}$, respectively. Therefore, the ML estimators of $\mathbf{H}_{11}$, $\mathbf{H}_{12}$, $\mathbf{H}_{21}$, $\mathbf{H}_{22}$, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are obtainable from Lemma 2.1. If $\mathbf{H}_{ij}$ is written in terms of $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}$ and $\boldsymbol{\lambda}$ defined in Section 3.2, we get

$$
\begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} = \begin{pmatrix} 2\boldsymbol{\mu} + \boldsymbol{\lambda} & 2\boldsymbol{\mu} - \boldsymbol{\lambda} \\ 2\boldsymbol{\tau} + 2\boldsymbol{\pi} - \boldsymbol{\lambda} & -2\boldsymbol{\tau} + 2\boldsymbol{\pi} + \boldsymbol{\lambda} \end{pmatrix}. \qquad (13)
$$

Let $\mathbf{v}_1 = (\boldsymbol{\mu}' : \boldsymbol{\lambda}')'$, $\mathbf{v}_2 = (\boldsymbol{\pi}' : \boldsymbol{\theta}')'$, $\mathbf{w}_1 = (\mathbf{H}_{11}' : \mathbf{H}_{12}')'$ and $\mathbf{w}_2 = (\mathbf{H}_{21}' : \mathbf{H}_{22}')'$, where $\boldsymbol{\theta} = 2\boldsymbol{\tau} - \boldsymbol{\lambda}$. The relation in (13) represents a linear transformation, L_1: $\mathbb{R}^{2q} \to \mathbb{R}^{2q}$, given by $L_1(\mathbf{v}_i) = \mathbf{M}_1\mathbf{v}_i$, where $\mathbf{M}_1$: $2q \times 2q$ equals

$$
\mathbf{M}_1 = \begin{pmatrix} 2\mathbf{I} & \mathbf{I} \\ 2\mathbf{I} & -\mathbf{I} \end{pmatrix}.
$$

L_1 is a one-to-one linear transformation, where the inverse of $\mathbf{M}_1$ equals

$$
\mathbf{M}_1^{-1} = \frac{1}{4} \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ 2\mathbf{I} & -2\mathbf{I} \end{pmatrix}.
$$

Let $\widehat{\mathbf{w}_i}$ be the ML estimators of $\mathbf{w}_i$, $i = 1, 2$. Then $L_1^{-1}(\widehat{\mathbf{w}_i}) = \widehat{L_1^{-1}(\mathbf{w}_i)} = \widehat{\mathbf{M}_1^{-1}\mathbf{w}_i} = \widehat{\mathbf{v}}_i$, $i = 1, 2$. Thus, the ML estimators of $\boldsymbol{\mu}$, $\boldsymbol{\lambda}$, $\boldsymbol{\pi}$ and $\boldsymbol{\theta}$ equal

$$\begin{aligned}
\widehat{\boldsymbol{\mu}} &= \frac{1}{4}(\widehat{\mathbf{H}}_{11} + \widehat{\mathbf{H}}_{12}),\\
\widehat{\boldsymbol{\lambda}} &= \frac{1}{2}(\widehat{\mathbf{H}}_{11} - \widehat{\mathbf{H}}_{12}),\\
\widehat{\boldsymbol{\pi}} &= \frac{1}{4}(\widehat{\mathbf{H}}_{21} + \widehat{\mathbf{H}}_{22}),\\
\widehat{\boldsymbol{\theta}} &= \frac{1}{2}(\widehat{\mathbf{H}}_{21} - \widehat{\mathbf{H}}_{22}).
\end{aligned}$$

The derivation of the ML estimator of $\boldsymbol{\tau}$ remains. Let $\mathbf{v}_3 = (\boldsymbol{\mu}' : \boldsymbol{\lambda}' : \boldsymbol{\pi}' : \boldsymbol{\tau}')'$ and $\mathbf{w}_3 = (\boldsymbol{\mu}' : \boldsymbol{\lambda}' : \boldsymbol{\pi}' : \boldsymbol{\theta}')'$. Let L_2: $\mathbb{R}^{4q} \to \mathbb{R}^{4q}$ be the linear transformation given by $L_2(\mathbf{x}) = \mathbf{M}_2\mathbf{x}$, where $\mathbf{M}_2$: $4q \times 4q$ equals

$$\mathbf{M}_2 = \begin{pmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & 2\mathbf{I} \end{pmatrix}.$$

Then $L_2(\mathbf{v}_3) = \mathbf{M}_2\mathbf{v}_3 = \mathbf{w}_3$ and L_2 is an one-to-one linear transformation, where the inverse of $\mathbf{M}_2$ equals

$$\mathbf{M}_2^{-1} = \frac{1}{2}\begin{pmatrix} 2\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & 2\mathbf{I} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & 2\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Therefore, $L_2^{-1}(\widehat{\mathbf{w}}_3) = \widehat{L_2^{-1}(\mathbf{w}_3)} = \widehat{\mathbf{M}_2^{-1}\mathbf{w}_3} = \widehat{\mathbf{v}}_3$, which means that the ML estimator of $\boldsymbol{\tau}$ equals

$$\widehat{\boldsymbol{\tau}} = \frac{1}{2}(\widehat{\boldsymbol{\lambda}} + \widehat{\boldsymbol{\theta}}) = \frac{1}{4}(\widehat{\mathbf{H}}_{11} - \widehat{\mathbf{H}}_{12} + \widehat{\mathbf{H}}_{21} - \widehat{\mathbf{H}}_{22}).$$

Alternatively, since $\boldsymbol{\tau} = \frac{1}{2}(\boldsymbol{\lambda} + \boldsymbol{\theta})$, it follows immediately that the maximum likelihood estimator of $\boldsymbol{\tau}$ must equal $\frac{1}{2}(\widehat{\boldsymbol{\lambda}} + \widehat{\boldsymbol{\theta}})$. In the same way, the ML estimators of $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{12}$ are obtained.

(ii) If $\rho(\mathbf{A}_1) = q$, then, since $\rho(\mathbf{C}) = 2$, the uniqueness follows immediately from Lemma 2.2. □

If $\boldsymbol{\lambda} = 0$, the estimator of $\boldsymbol{\tau}$ follows easily due to the independence of the column vectors $\mathbf{X}_{1j} + \mathbf{X}_{2j}$ and $\mathbf{X}_{1j} - \mathbf{X}_{2j}$.

Corollary 3.4 *Suppose* $\boldsymbol{\lambda} = \mathbf{0}$ *in the model defined by Theorem 3.3. Then the maximum likelihood estimator of* $\boldsymbol{\tau}$ *is equal to*

$$\widehat{\boldsymbol{\tau}} = \frac{1}{2}\widehat{\boldsymbol{\theta}} = \frac{1}{4}(\widehat{\mathbf{H}}_{21} - \widehat{\mathbf{H}}_{22}).$$

3.5 A two-sequence four-period design

In this section the two-sequence four-period ABAB:BABA cross-over design is considered. As before, it is assumed that a time series of p measurements is collected within periods at the same time points for each subject, and that the vectors of observations from subjects are independently multivariate normally distributed. The same notations as in Section 3.2 are used as much as possible. Let $\mathbf{X}$: $4p \times n$ be an observation matrix, $\mathbf{A}$: $4p \times 4q$ be a within-subject design matrix, $\mathbf{B}$: $4q \times 2$ be an unknown parameter matrix, $\mathbf{C}$: $2 \times (n_1 + n_2)$ be a between-sequence design matrix and $\boldsymbol{\Sigma}$: $4p \times 4p$ be an unknown p.d. dispersion matrix. Moreover,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{1}'_{n_1} & \mathbf{0}' \\ \mathbf{0}' & \mathbf{1}'_{n_2} \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \\ \mathbf{B}_{31} & \mathbf{B}_{32} \\ \mathbf{B}_{41} & \mathbf{B}_{42} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_1 + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_1 & \boldsymbol{\mu}_2 + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_1 \\ \boldsymbol{\mu}_1 + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_2 + \boldsymbol{\lambda}_1 & \boldsymbol{\mu}_2 + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_2 + \boldsymbol{\lambda}_2 \\ \boldsymbol{\mu}_1 + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_3 + \boldsymbol{\lambda}_2 + \boldsymbol{\theta}_1 & \boldsymbol{\mu}_2 + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_3 + \boldsymbol{\lambda}_1 + \boldsymbol{\theta}_2 \\ \boldsymbol{\mu}_1 + \boldsymbol{\tau}_2 + \boldsymbol{\pi}_4 + \boldsymbol{\lambda}_1 + \boldsymbol{\theta}_2 & \boldsymbol{\mu}_2 + \boldsymbol{\tau}_1 + \boldsymbol{\pi}_4 + \boldsymbol{\lambda}_2 + \boldsymbol{\theta}_1 \end{pmatrix},$$

where $\mathbf{A}_1$: $p \times q$ is defined as in Section 3.2. Now the model becomes the usual GCM1. One difference with the two-period model is that we now have second-order carry-over effects, $\boldsymbol{\theta}_i$, $i = 1, 2$. The ML estimators of $\boldsymbol{\mu}_i$, $\boldsymbol{\tau}_i$, $\boldsymbol{\pi}_i$, $\boldsymbol{\lambda}_i$ and $\boldsymbol{\theta}_i$, $i = 1, 2$, can be derived in the same way as in Section 3.3. The results are summarized in the next theorem.

Theorem 3.5 *Suppose that we have a repeated measures two-sequence four-period ABAB:BABA cross-over design with* n_1 *and* n_2 *subjects in treatment sequence one and two, respectively. Suppose that a response variable is measured at time points* $t_1, \ldots, t_p$ *within periods for each subject, and that the* $4p$*-vectors of outcomes from subjects are independent observations from a multivariate normal distribution with an unknown positive definite dispersion matrix,* $\boldsymbol{\Sigma}$*. Suppose* $n = n_1 + n_2 \geq 4p + 2$*. Let further* $\mathbf{A}$*,* $\mathbf{A}_1$*,* $\mathbf{B}$ *and* $\mathbf{C}$ *be defined as above.*

(i) *If the constraints*

$$\boldsymbol{\tau}_1 + \boldsymbol{\tau}_2 = \mathbf{0}, \tag{14}$$

$$\boldsymbol{\pi}_1 + \boldsymbol{\pi}_2 + \boldsymbol{\pi}_3 + \boldsymbol{\pi}_4 = \mathbf{0}, \tag{15}$$

$$\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2 = \mathbf{0}, \tag{16}$$

$$\boldsymbol{\theta}_1 + \boldsymbol{\theta}_2 = \mathbf{0} \tag{17}$$

are applied to the parameters and they are renamed as $\boldsymbol{\tau} = \boldsymbol{\tau}_1$, $\boldsymbol{\lambda} = \boldsymbol{\lambda}_1$ *and* $\boldsymbol{\theta} = \boldsymbol{\theta}_1$, *then the maximum likelihood estimators of the parameters equal*

$$\widehat{\boldsymbol{\mu}}_1 = \frac{1}{8}(\mathbf{I} : \mathbf{I} : 3\mathbf{I} : 3\mathbf{I} : \mathbf{I} : \mathbf{I} : -\mathbf{I} : -\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \frac{1}{8}(\widehat{\mathbf{B}}_{11} + \widehat{\mathbf{B}}_{21} + 3\widehat{\mathbf{B}}_{31} + 3\widehat{\mathbf{B}}_{41} + \widehat{\mathbf{B}}_{12} + \widehat{\mathbf{B}}_{22} - \widehat{\mathbf{B}}_{32} - \widehat{\mathbf{B}}_{42}), \tag{18}$$

$$\widehat{\boldsymbol{\mu}}_2 = \frac{1}{8}(\mathbf{I} : \mathbf{I} : -\mathbf{I} : -\mathbf{I} : \mathbf{I} : \mathbf{I} : 3\mathbf{I} : 3\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \frac{1}{8}(\widehat{\mathbf{B}}_{11} + \widehat{\mathbf{B}}_{21} - \widehat{\mathbf{B}}_{31} - \widehat{\mathbf{B}}_{41} + \widehat{\mathbf{B}}_{12} + \widehat{\mathbf{B}}_{22} + 3\widehat{\mathbf{B}}_{32} + 3\widehat{\mathbf{B}}_{42}), \tag{19}$$

$$\widehat{\boldsymbol{\tau}} = \frac{1}{8}(4\mathbf{I} : \mathbf{0} : -2\mathbf{I} : -2\mathbf{I} : -4\mathbf{I} : \mathbf{0} : 2\mathbf{I} : 2\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \frac{1}{8}(4\widehat{\mathbf{B}}_{11} - 2\widehat{\mathbf{B}}_{31} - 2\widehat{\mathbf{B}}_{41} - 4\widehat{\mathbf{B}}_{12} + 2\widehat{\mathbf{B}}_{32} + 2\widehat{\mathbf{B}}_{42}), \tag{20}$$

$$\widehat{\boldsymbol{\pi}}_1 = \frac{1}{8}(3\mathbf{I} : -\mathbf{I} : -\mathbf{I} : -\mathbf{I} : 3\mathbf{I} : -\mathbf{I} : -\mathbf{I} : -\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \frac{1}{8}(3\widehat{\mathbf{B}}_{11} - \widehat{\mathbf{B}}_{21} - \widehat{\mathbf{B}}_{31} - \widehat{\mathbf{B}}_{41} + 3\widehat{\mathbf{B}}_{12} - \widehat{\mathbf{B}}_{22} - \widehat{\mathbf{B}}_{32} - \widehat{\mathbf{B}}_{42}), \tag{21}$$

$$\widehat{\boldsymbol{\pi}}_2 = \frac{1}{8}(-\mathbf{I} : 3\mathbf{I} : -\mathbf{I} : -\mathbf{I} : -\mathbf{I} : 3\mathbf{I} : -\mathbf{I} : -\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \frac{1}{8}(3\widehat{\mathbf{B}}_{11} + 3\widehat{\mathbf{B}}_{21} - \widehat{\mathbf{B}}_{31} - \widehat{\mathbf{B}}_{41} - \widehat{\mathbf{B}}_{12} + 3\widehat{\mathbf{B}}_{22} - \widehat{\mathbf{B}}_{32} - \widehat{\mathbf{B}}_{42}), \tag{22}$$

$$\widehat{\boldsymbol{\pi}}_3 = \frac{1}{8}(-\mathbf{I} : -\mathbf{I} : 3\mathbf{I} : -\mathbf{I} : -\mathbf{I} : -\mathbf{I} : 3\mathbf{I} : -\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \frac{1}{8}(-\widehat{\mathbf{B}}_{11} - \widehat{\mathbf{B}}_{21} + 3\widehat{\mathbf{B}}_{31} - \widehat{\mathbf{B}}_{41} - \widehat{\mathbf{B}}_{12} - \widehat{\mathbf{B}}_{22} + 3\widehat{\mathbf{B}}_{32} - \widehat{\mathbf{B}}_{42}), \tag{23}$$

$$\widehat{\boldsymbol{\pi}}_4 = \mathbf{1} - \widehat{\boldsymbol{\pi}}_1 - \widehat{\boldsymbol{\pi}}_2 - \widehat{\boldsymbol{\pi}}_3 \quad (\mathbf{1} \textit{ is a vector of ones}), \tag{24}$$

$$\widehat{\boldsymbol{\lambda}} = \frac{1}{8}(4\mathbf{I} : 4\mathbf{I} : -4\mathbf{I} : -4\mathbf{I} : -4\mathbf{I} : -4\mathbf{I} : 4\mathbf{I} : 4\mathbf{I})\mathrm{vec}\widehat{\mathbf{B}}$$

$$= \tfrac{1}{8}(4\widehat{\mathbf{B}}_{11} + 4\widehat{\mathbf{B}}_{21} - 4\widehat{\mathbf{B}}_{31} - 4\widehat{\mathbf{B}}_{41} - 4\widehat{\mathbf{B}}_{12} - 4\widehat{\mathbf{B}}_{22} + 4\widehat{\mathbf{B}}_{32} + 4\widehat{\mathbf{B}}_{42}), \tag{25}$$

$$\begin{aligned}
\widehat{\boldsymbol{\theta}} &= \frac{1}{8}(\mathbf{0} : 4\mathbf{I} : \mathbf{0} : -4\mathbf{I} : \mathbf{0} : -4\mathbf{I} : \mathbf{0} : 4\mathbf{I})\text{vec}\widehat{\mathbf{B}} \\
&= \frac{1}{8}(4\widehat{\mathbf{B}}_{21} - 4\widehat{\mathbf{B}}_{41} - 4\widehat{\mathbf{B}}_{22} + 4\widehat{\mathbf{B}}_{42}), \qquad (26) \\
\widehat{\boldsymbol{\Sigma}} &= \frac{1}{n}(\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C})(\mathbf{X} - \mathbf{A}\widehat{\mathbf{B}}\mathbf{C})', \qquad (27)
\end{aligned}$$

where $\widehat{\mathbf{B}}$ *is presented in Lemma 2.1.*

(ii) *The ML estimators of* $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}_1$, $\boldsymbol{\pi}_2$, $\boldsymbol{\pi}_3$, $\boldsymbol{\pi}_4$, $\boldsymbol{\lambda}$ *and* $\boldsymbol{\theta}$ *are unique if and only if* $\rho(\mathbf{A}_1) = q$.

Proof Since the proof is similar to the proof of Theorem 3.1 some details will be left out.

(i) The model satisfies all conditions of a GCM1 (see Definition 2.1) and the ML estimators of $\mathbf{B}$ and $\boldsymbol{\Sigma}$ are given by Lemma 2.1. After applying the constraints given by (14)-(17), there will be an one-to-one linear transformation between $\mathbf{v} = (\boldsymbol{\mu}_1' : \boldsymbol{\mu}_2' : \boldsymbol{\tau}' : \boldsymbol{\pi}_1' : \boldsymbol{\pi}_2' : \boldsymbol{\pi}_3' : \boldsymbol{\lambda}' : \boldsymbol{\theta}')'$ and $\mathbf{w} = \text{vec}\mathbf{B}$, say $L(\mathbf{v}) = \mathbf{M}\mathbf{v} = \mathbf{w}$, where $\mathbf{M}$: $8q \times 8q$ and $\mathbf{M}^{-1}$ are given by

$$\mathbf{M} = \begin{pmatrix}
\mathbf{I} & \mathbf{0} & \mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{I} & \mathbf{0} & -\mathbf{I} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\
\mathbf{I} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{I} & -\mathbf{I} & \mathbf{I} \\
\mathbf{I} & \mathbf{0} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & \mathbf{I} & -\mathbf{I} \\
\mathbf{0} & \mathbf{I} & -\mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{I} & \mathbf{I} & \mathbf{0} & \mathbf{I} & \mathbf{0} & -\mathbf{I} & \mathbf{0} \\
\mathbf{0} & \mathbf{I} & -\mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{I} & -\mathbf{I} \\
\mathbf{0} & \mathbf{I} & \mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & \mathbf{I}
\end{pmatrix}$$

and

$$\mathbf{M}^{-1} = \frac{1}{8}\begin{pmatrix}
\mathbf{I} & \mathbf{I} & 3\mathbf{I} & 3\mathbf{I} & \mathbf{I} & \mathbf{I} & -\mathbf{I} & -\mathbf{I} \\
\mathbf{I} & \mathbf{I} & -\mathbf{I} & -\mathbf{I} & \mathbf{I} & \mathbf{I} & 3\mathbf{I} & 3\mathbf{I} \\
4\mathbf{I} & \mathbf{0} & -2\mathbf{I} & -2\mathbf{I} & -4\mathbf{I} & \mathbf{0} & 2\mathbf{I} & 2\mathbf{I} \\
3\mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & 3\mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} \\
-\mathbf{I} & 3\mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & 3\mathbf{I} & -\mathbf{I} & -\mathbf{I} \\
-\mathbf{I} & -\mathbf{I} & 3\mathbf{I} & -\mathbf{I} & -\mathbf{I} & -\mathbf{I} & 3\mathbf{I} & -\mathbf{I} \\
4\mathbf{I} & 4\mathbf{I} & -4\mathbf{I} & -4\mathbf{I} & -4\mathbf{I} & -4\mathbf{I} & 4\mathbf{I} & 4\mathbf{I} \\
\mathbf{0} & 4\mathbf{I} & \mathbf{0} & -4\mathbf{I} & \mathbf{0} & -4\mathbf{I} & \mathbf{0} & 4\mathbf{I}
\end{pmatrix}.$$

Thus, the ML estimator of $\mathbf{v}$ is given by $\widehat{\mathbf{v}} = L^{-1}(\widehat{\mathbf{w}}) = \mathbf{M}^{-1}\widehat{\mathbf{w}}$.

(ii) Since $\rho(\mathbf{A}_1) = q$ if and only if $\rho(\mathbf{A}) = 4q$, the result follows from Lemma 2.1 (ii). □

Next, a model where the second order effect parameter $\boldsymbol{\theta}$ in Theorem 3.5 equals $\mathbf{0}$ is considered.

Theorem 3.6 *Suppose that the same conditions as in Theorem 3.5 hold. In addition, suppose* $\boldsymbol{\theta} = \mathbf{0}$. *Then the maximum likelihood estimators of the parameters* $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{\tau}$, $\boldsymbol{\pi}_1$, $\boldsymbol{\pi}_2$, $\boldsymbol{\pi}_3$, $\boldsymbol{\pi}_4$, $\boldsymbol{\lambda}$ *are given by (18)-(25) and* $\boldsymbol{\Sigma}$ *by (27), with* $\widehat{\mathbf{B}}$ *presented in Lemma 2.2.*

Proof Let $\mathbf{D}$: $q \times 4q$ and $\mathbf{F}$: 2×1 be defined by

$$\mathbf{D} = \begin{pmatrix} \mathbf{0}: & \mathbf{I}: & \mathbf{0}: & -\mathbf{I} \end{pmatrix} \quad \text{and} \quad \mathbf{F} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Then $\mathbf{DBF} = 2\boldsymbol{\theta}$ and consequently $\mathbf{DBF} = \mathbf{0}$ if and only if $\boldsymbol{\theta} = \mathbf{0}$. ML estimators of the parameters are now obtainable via Lemma 2.2 and 2.3. □

3.6 A numerical example

To illustrate the results in Sections 3.1-3.5, a study published by Ciminera and Wolfe (1953) (re-printed by Kenward and Jones (1987) and re-analyzed by Putt and Chinchilli (1999)), is analyzed. The aim of this is to demonstrate the models and the calculations of the maximum likelihood estimates without making any formal statistical inference. It is a study of the blood sugar level (mg %) in female rabbits for two different insulin mixtures called A and B. The design is a two-sequence four-period ABAB:BABA cross-over design, where the the blood sugar level was measured at five successive times ($t = 0, 1.5, 3, 3, 4.5$ h) after injection. Consequently, the results in Section 3.5 are applied. In addition, in order to apply the results in Sections 3.2-3.4 for an AB:BA design, an analysis is made of the two first periods of the study. As observed by Putt and Chinchilli (1999), the data suggest the blood sugar level to be modelled with a second-order polynomial function of time. Therefore, the within-subject design matrix $\mathbf{A}_1$ equals

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1.5 & 1.5^2 \\ 1 & 3 & 3^2 \\ 1 & 4.5 & 4.5^2 \\ 1 & 6 & 6^2 \end{pmatrix}.$$

There were 11 rabbits included in each treatment sequence, which means the between-subject matrix $\mathbf{C}$ equals

$$\mathbf{C} = \begin{pmatrix} \mathbf{1}'_{11} & \mathbf{0}'_{11} \\ \mathbf{0}'_{11} & \mathbf{1}'_{11} \end{pmatrix}.$$

The following models are evaluated:

Model 1a: The model defined in Sections 3.2 – 3.3 for an AB:BA cross-over design, where a parameter for a carry-over effect is included. Theorem 3.1 is used to estimate the parameters. The results are presented in Table 1.

Model 1b: The model defined in Sections 3.2-3.3 for an AB:BA cross-over design, where no carry-over is included. Theorem 3.2 is used to estimate the parameters. The results are presented in Table 1.

Model 2: The model defined in Section 3.4 for an AB:BA cross-over design. The main model includes a parameter for a carry-over effect but the treatment effect is also estimated under the assumption of no carry-over effect. Theorem 3.3 and Corollary 3.4 are used to estimate the parameters. The results are presented in Table 2.

Model 3a: The model defined in Section 3.5.1 for an ABAB:BABA cross-over design, where parameters for a first- and second-order carry-over effects are included. Theorem 3.5 is used to estimate the parameters. The results are presented in Table 3.

Model 3b: The model defined in Section 3.5.1 for an ABAB:BABA cross-over design, in which the second-order carry-over effect is omitted. Theorem 3.6 is used to estimate the parameters. The results are presented in Table 3.

For all models the estimated standard deviations of the estimated parameters have been calculated by help of Lemma 2.1 (iii) and Lemma 2.3 (iii), where $\mathbf{\Sigma}$ has been replaced by its ML estimator, and the fact that the dispersion matrix of $\mathbf{M}\widehat{\mathbf{B}}$, where $\mathbf{M}$ is a known matrix of proper size, equals $D[\mathbf{M}\widehat{\mathbf{B}}] = \mathbf{M}D[\widehat{\mathbf{B}}]\mathbf{M}'$.

Since no hypothesis testing is made, no formal inference about the parameter estimates is made. However, all models seem to fit data well with respect to the mean response profile. The conclusion by Putt and Chinchilli (1999), and Cinemera and Wolfe (1953), was that treatment B resulted in a slightly larger linear slope than treatment A. This in agreement with the results of Model 1b, Model 2 ($\widehat{\boldsymbol{\tau}}|\boldsymbol{\lambda} = \mathbf{0}$) and Model 3a-3b, and the magnitudes of the treatment difference (A-B) are also in the same range. Observe that the constraints on the parameters (see (2) and (14)) imply that the treatment difference (A-B) equals $2\boldsymbol{\tau}$.

TABLE 1. Maximum likelihood estimates of the parameters for Model 1a and 1b, by use of Theorem 3.1 and 3.2, respectively. The standard deviations of the estimates are given in brackets.

Model	Coefficient	$\widehat{\boldsymbol{\mu}}$	$\widehat{\boldsymbol{\tau}}$	$\widehat{\boldsymbol{\pi}}$	$\widehat{\boldsymbol{\lambda}}$
	Intercept	87.78	0.30	1.49	-0.69
		(1.54)	(2.03)	(1.06)	(1.54)
Model 1a	Linear	-38.32	2.23	-0.93	7.69
		(2.27)	(2.57)	(0.79)	(2.27)
	Quadratic	6.29	-0.52	0.08	-1.37
		(0.38)	(0.46)	(0.02)	(0.38)
	Intercept	87.89	1.32	1.71	.
		(2.54)	(0.82)	(1.12)	.
Model 1b	Linear	-39.58	-2.17	-1.11	.
		(3.97)	(0.53)	(0.97)	.
	Quadratic	6.51	0.25	0.11	.
		(0.68)	(0.12)	(0.20)	.

TABLE 2. Maximum likelihood estimates of the parameters for Model 2, by use of Theorem 3.3 and Corollary 3.4. The standard deviations of the estimates are given in brackets.

Model	Coefficient	$\widehat{\boldsymbol{\mu}}$	$\widehat{\boldsymbol{\tau}}$	$\widehat{\boldsymbol{\pi}}$	$\widehat{\boldsymbol{\lambda}}$	$\widehat{\boldsymbol{\tau}}\vert\boldsymbol{\lambda}=\mathbf{0}$
	Intercept	88.49	0.58	0.82	-0.49	0.82
		(1.41)	(1.71)	(0.96)	(2.82)	(0.96)
Model 2	Linear	-37.28	3.28	0.93	9.93	-1.69
		(2.12)	(2.24)	(0.74)	(4.23)	(0.73)
	Quadratic	6.02	-0.75	-0.26	-1.86	0.18
		(0.36)	(0.39)	(0.15)	(0.72)	(0.15)

TABLE 3. Maximum likelihood estimates of the parameters for Model 3a and 3b, by use of Theorem 3.6 and 3.7, respectively. The standard deviations of the estimates are given in brackets.

Model	Coefficient	$\widehat{\boldsymbol{\mu}}_1$	$\widehat{\boldsymbol{\mu}}_2$	$\widehat{\boldsymbol{\tau}}$	$\widehat{\boldsymbol{\pi}}_1$	$\widehat{\boldsymbol{\pi}}_2$	$\widehat{\boldsymbol{\pi}}_3$	$\widehat{\boldsymbol{\lambda}}$	$\widehat{\boldsymbol{\theta}}$
	Intercept	92.50	86.88	-2.00	-2.77	-3.62	-4.24	-5.40	-0.82
		(3.09)	(3.09)	(2.46)	(1.66)	(1.17)	(2.38)	(3.74)	(3.50)
Model 3a	Linear	-35.96	-45.82	-1.77	-7.73	-3.29	1.30	-0.83	0.74
		(2.64)	(2.64)	(2.95)	(1.78)	(1.34)	(2.53)	(5.11)	(4.46)
	Quadratic	5.84	7.36	0.09	1.45	0.83	-0.34	-0.11	-0.28
		(0.43)	(0.43)	(0.58)	(0.37)	(0.24)	(0.37)	(0.92)	(0.21)
	Intercept	91.40	87.84	-1.07	-2.71	-3.63	-4.21	-4.34	.
		(2.12)	(2.12)	(1.44)	(1.35)	(1.16)	(1.70)	(2.19)	.
Model 3b	Linear	-36.39	-45.42	-2.58	-7.77	-3.31	1.32	-2.23	.
		(1.93)	(1.93)	(1.42)	(1.57)	(1.32)	(1.73)	(2.25)	.
	Quadratic	5.90	7.32	0.23	1.45	0.84	-0.34	0.23	.
		(0.32)	(0.32)	(0.29)	(0.31)	(0.24)	(0.24)	(0.39)	.

4 Conclusion and discussion

Models for the repeated measures AB:BA and ABAB:BABA cross-over design have been defined in terms of the GCM1, and maximum likelihood estimators of the parameters have been derived by utilizing the theory for multivariate linear normal models. The models differ from the standard situation in that the main parameters, which we call the cross-over parameters, are linear combinations $\mathbf{DBF}$ of $\mathbf{B}$, for some matrices $\mathbf{D}$ and $\mathbf{F}$, whereas the matrix $\mathbf{B}$ generally is not of interest. The number of cross-over parameters is greater than the number of elements of $\mathbf{B}$. However, by placing certain natural additional restrictions (see (2)-(4) and (14)-(17)) on the cross-over parameters, the cross-over parameters and the elements of $\mathbf{B}$ are related according to a one-to-one transformation. The additional restrictions of the parameters will not change the estimator of any treatment contrast. Other types of restrictions than those suggested may also work, only the interpretation of the parameters will change. To illustrate numerically, the maximum likelihood estimates of the parameters for a study published by Cimenera and Wolfe (1953) have been calculated. The results are in agreement with the analysis made by Putt and Chinchilli (1999) using a mixed effect models, although no formal statistical inference of the parameters has been made.

The advantage of the approach is that explicit expressions for the estimators of the parameters are obtained. Explicit expression for an estimator is advantageous since then, for example, the distribution of the estimator can be approximated and properties of the distributions examined. A disadvantage of the model is that the number of parameters for estimating the error structure is greater than in a mixed effect model, in which a certain structure of the error matrix could reduce the number parameters (see Putt and Chinchilli, 1999). However, it is also possible in the GCM1 and GCM2 to assume certain structures in the dispersion matrix. This is demonstrated in Section 3.4 for the AB:BA design, where $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}'_{12}$ and $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22}$ are supposed to hold.

There remain several questions about the methodology that could be of interest in future work. For example, hypothesis testing of the parameters in the model and other types of structures of the dispersion matrix than given in Section 3.4 are areas to be resolved. Moreover, further experience when the models are applied to real data is needed.

References

Banken, L. (1984). Eine verallgemeinerung des GMANOVA modells. Dissertation, University of Trier, Trier, Germany.

Chinchilli, V.M. and Esinhart, J.D. (1996). Design and analysis of intra-subject variability in cross-over experiments. *Statistics in Medicine* **15**, 1619-1634.

Ciminera, J.L. and Wolfe, E.K. (1953). An example of the use of extended cross-over design in the comparison of NPH insulin mixtures. *Biometrics* **9**, 431-446.

Gleser, L.J. and Olkin, I. (1970). Linear models in multivariate analysis. *Essays in Probability and Statistics*, 267-292, University of North Carolina Press, Chapel Hill, North Carolina.

Jones, B. and Kenward, M.G. (1987). *Design and Analysis of Cross-Over Trials,* Chapman and Hall, London.

Kshirsagar, A.M. and Smith, W.B. (1995). *Growth Curves*, Marcel Dekker, New York.

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

Putt, M. and Chinchilli, V.M. (1999). A mixed effects model for the analysis of repeated measures cross-over studies. *Statistics in Medicine* **18**, 3037-3058.

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

von Rosen, D. (1989). Maximum likelihood estimators in multivariate linear normal models. *J. Multivar. Analysis* **31**, 187-200.

von Rosen, D. (1990). Moments for a multivariate linear model with an application to the growth curve model. *J. Multivar. Analysis* **35**, 243-259.

von Rosen, D. (1991). The growth curve model: A review. *Comm. Statist. Theory. Methods* **20**, 2791-2822.

von Rosen, D. (1995). Uniqueness conditions for maximum likelihood estimators in a multivariate linear normal model. *J. Statist. Plann. Inference* **36**, 269-275.

Senn, S. (1993). *Cross-Over Trials in Clinical Research*, Wiley, New York.

Åsenblad, N. (2001). Multivariate linear normal models for the analysis of cross-over designs. Licentiate thesis, Department of Biometry and Informatics, Swedish University of Agricultural Sciences, Uppsala, Sweden.