Electronic Journal of Statistics
Vol. 7 (2013) 244–263
ISSN: 1935-7524
DOI: 10.1214/13-EJS770

Identifiability of linear mixed effects models

Wei Wang

Center for Outcomes Research, Children's Hospital of Philadelphia, 3535 Market Street, Suite 1029, Philadelphia, PA, 19104, USA
e-mail: wangw@email.chop.edu

Abstract: In linear mixed effects models, the covariance matrix of the response is modeled as the sum of two matrices: the product of the covariance matrix of the random effects with the associated design matrix, and the covariance matrix of the residual error. Building a linear mixed model usually involves selection of the parametrized covariance matrix structures for the random effects and the residual error. However, even if the covariance matrix of the response is not over-parametrized, some specifications of covariance structures can result in the non-identifiability of parameters. When fitting such models, software may or may not indicate a problem with model identifiability. Consequently, it is useful to have a way to check if a model is identifiable which does not rely on the software output. We derive conditions for identifiability of the covariance parameters of the response and study commonly used covariance structures. The derived conditions only rely on the covariance structures being used and properties of the design matrix associated with the random effects and are easy to check.

AMS 2000 subject classifications: Primary 60K35.
Keywords and phrases: Covariance matrix structures, identifiability, linear mixed effects models, random effects, variance components.

Received October 2012.

Contents

1. Introduction

Study designs in a variety of disciplines such as agriculture, biology, medicine, physics, and social sciences may lead to data sets of a clustered structure, or a longitudinal or repeated-measures structure. A linear mixed effects model (LME) may be well suited for the analysis of these types of data. LME has been developed, studied and applied extensively by many researchers, for example, [2, 5, 6, 7, 8, 12, 15, 17].

In LME, the response is modeled as a sum of three terms: the design matrix associated with the fixed effects right multiplied by an unknown coefficient vector, the random effects left multiplied by the associated design matrix, and a residual error. The mean of the response is parameterized as the first term in the above sum as the random effects and the residual error are usually assumed mean zero. So long as that design matrix is of full column rank, there is no other coefficient vector giving the same mean. The covariance of the response is determined through the covariances of the random effects and of the residual error. Both the random effects and the residual error are assumed to have certain distributions with covariance matrix structures to be specified. The associated unknown parameters are usually constrained in certain spaces and are to be estimated. For a comprehensive list of available structures and the expressions, see for instance [13, 19].

The most commonly used types of structures include unstructured (UN), compound symmetry (CS), variance components (VC), multiple of a known covariance matrix (MK) or multiple of the identity matrix (MI), and one deriving from time series such as MA(1). For instance, [11] consider models where random effects are MI structured. [14] consider a variety of combination of structures, including UN random effects and UN error. [19] provides a comprehensive comparison of various combinations of structures. Choices of the various covariance structures are implemented in statistical software such as R, SAS, SPSS and STATA. Detailed documentation has been provided by [10, 13, 16].

Building an LME often involves selection of covariance structures for the random effects and the error (see for instance [1, 18, 19]). The covariance of the response is then modeled as a sum involving the two covariance matrices. In the modeling, some choices of covariance structures can lead to parameters not identifiable even if the response covariance is not over-parametrized. For instance, suppose the design matrix associated with the random effects has the form $\mathbf{Z} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. Let $\mathbf{I}_2$ be a 2×2 identity matrix. Suppose one specifies an MI structure for both the random effects and the residual error as $\sigma_u^2 \mathbf{I}_2$ and $\sigma^2 \mathbf{I}_2$ respectively. Elements of the parameter vector (σ_u^2, σ^2) are both positive and are to be estimated. Let $\mathbf{y}$ be the response vector with covariance matrix $\mathbf{\Sigma}_y$. Then $\forall\ \sigma_u^2 > 0$ and $\sigma^2 > 0$, there exists a distinct parameter vector $(\sigma_u^{*2}, \sigma^{*2})$ producing the same $\mathbf{\Sigma}_y$. Let r be a real number such that $0 < r < 1 + 2\sigma_u^2/\sigma^2$ and $r \neq 1$. Define $\sigma^{*2} = r\sigma^2$ and $\sigma_u^{*2} = \sigma_u^2 + (1-r)\sigma^2/2$. It is clear $\sigma_u^{*2} \neq \sigma_u^2,\ \sigma_u^{*2} > 0$, and $\sigma^{*2} \neq \sigma^2,\ \sigma^{*2} > 0$, but $\mathbf{\Sigma}_y = (2\sigma_u^2 + \sigma^2)\mathbf{I}_2 = (2\sigma_u^{*2} + \sigma^{*2})\mathbf{I}_2$. If the research interest relates to the covariance parameters, say the BLUP, we see that the BLUP can either be $\sigma_u^2 \mathbf{Z}\mathbf{\Sigma}_y^{-1}(\mathbf{y} - \mathrm{E}\mathbf{y})$ or $\sigma_u^{*2} \mathbf{Z}\mathbf{\Sigma}_y^{-1}(\mathbf{y} - \mathrm{E}\mathbf{y})$.

In principle, when a model is specified, its identifiability shall be checked before fitting [18]. Identifiability may also be viewed as a necessary property for the adequacy of a statistical model [2], and is closely related to parameter estimability [3]. Non-identifiability has been noticed for the LME's in [9] (revisited in Example 5.1), [18] (revisited in Example 5.3), and [13] where one of the random effects terms is totally confounded with the residual error. [9] observe non-identifiability by calculating the covariance matrix elements of the response directly. Direct calculation may provide a way to detect non-identifiability if the model is not overly complicated. The models in [14] and [19] (revisited in Example 5.2) are very similar to those in [9] and are, in fact, identifiable. When fitting non-identifiable models, software may produce error messages of non-convergence or failure of constructing confidence intervals on the parameters [18]. Another possible indication of non-identifiability is unreasonably wide confidence intervals of the standard deviations [13], or zero or abnormally large estimated standard errors of the parameters [18]. However, in Section 3, the simulations of fitting different non-identifiable models, we found about half of the times the software output did not look unreasonable. Consequently, it is useful to have a way of checking if a model is identifiable before fitting.

In this article, we study identifiability of the covariance parameters in LME, focusing on those models that are not over-parametrized. We derive conditions of identifiability and study commonly used covariance structures. The results do not rely on any distribution assumptions of the random effects or the residual error. One condition also generalizes a known result of checking identifiability of an LME in [2]. The rest of the article is organized as follows. In Section 2, we give the notation and setup of the study. Simulation studies are presented in Section 3. The derived conditions are presented in Section 4. In Section 5, we show identifiability or non-identifiability of models in [9, 14, 18, 19] using the results in Section 4. At last, we discuss applicability of the obtained results to verify identifiability for models with more than two random components. Proofs are presented in the appendix.

2. Notation and setup

Given a parametrized distribution, identifiability requires that identical parameter vectors produce the same distribution (or distinct parameter vectors produce different distributions) [2, 3]. As noted in [2], non-identifiability of the covariance parameters is a sufficient condition of distribution non-identifiability. If a random variable is normally distributed, identifiability of parameters in its first two moments is then equivalent to identifiability of the distribution. An LME explicitly formulates the distribution of the response. As noticed, in an LME, identifiability of the mean parameters just requires the design matrix associated with the fixed effects to be of full column rank. If the fixed effects design is not of full column rank, one can estimate functions of the parameters that are identifiable, or reparameterize the fixed effects to induce identifiability. For

simplicity, we always refer distribution parameters to those in the covariance matrices unless otherwise specified, and we do not distinguish identifiability of covariance parameters and identifiability of model parameters and use the two terms interchangeably.

Usually in an LME, covariance structures are specified with the constrained unknown parameters in a space to be estimated. The study of non-identifiability aims for any unknown covariance parameter values in the space. Given any parameter vector, we show existence of a distinct parameter vector in the same space producing the same model. Below the non-identifiability definition is in a form closely related to our discussion, followed by a review of the LME setup.

Definition 2.1. *Let* $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ *be a matrix parametrized by a vector* $\boldsymbol{\theta}$ *in a space* $\boldsymbol{\Theta} = \{\boldsymbol{\theta} : \ \boldsymbol{\Sigma}(\boldsymbol{\theta})$ *is symmetric and positive definite*$\}$*. The matrix* $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ *is not identifiable with respect to* $\boldsymbol{\theta}$ *on* $\boldsymbol{\Theta}$ *if* $\forall\ \boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\exists\ \boldsymbol{\theta}^* \in \boldsymbol{\Theta}$, *with* $\boldsymbol{\theta}^* \neq \boldsymbol{\theta}$, *such that* $\boldsymbol{\Sigma}(\boldsymbol{\theta}^*) = \boldsymbol{\Sigma}(\boldsymbol{\theta})$.

Let $\mathbf{y}$ be a response vector of length n and let $\mathbf{X}$ and $\mathbf{Z}$ be known, non-random design matrices where $\mathbf{Z}$ is of dimension $n \times q,\ n \geq q$. The linear mixed model is written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\epsilon},$$

$$\mathbf{u} \sim (\mathbf{0}, \boldsymbol{\Sigma}_u),\ \boldsymbol{\epsilon} \sim (\mathbf{0}, \boldsymbol{\Sigma}_\epsilon),\ \mathbf{u} \text{ independent of } \boldsymbol{\epsilon}. \tag{1}$$

Let $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_\epsilon$ be parametrized by $\boldsymbol{\theta}_u \in \boldsymbol{\Theta}_u$ and $\boldsymbol{\theta}_\epsilon \in \boldsymbol{\Theta}_\epsilon$ respectively. Unknown parameters in the model are then $(\boldsymbol{\beta}, \boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$, We study the parametrized covariance matrix of $\mathbf{y}$, $\boldsymbol{\Sigma}_y(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon) = \mathbf{Z}\boldsymbol{\Sigma}_u(\boldsymbol{\theta}_u)\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta}_\epsilon)$.

Since $\boldsymbol{\Sigma}_y(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$ is symmetric, the number of distinct elements is less than or equal to $n(n+1)/2$. If the total number of parameters in $(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$ exceeds $n(n+1)/2$, $\boldsymbol{\Sigma}_y(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$ is then over-parameterized and most likely not identifiable. Throughout the article, we make the following assumptions unless otherwise specified: parameter vectors $\boldsymbol{\beta}$, $\boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_\epsilon$ do not share common elements, $\boldsymbol{\Sigma}_u(\boldsymbol{\theta}_u)$ and $\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta}_\epsilon)$ are each identifiable with respect to its parameters, $\mathbf{Z}$ has full column rank. By the second assumption, $\boldsymbol{\Sigma}_u(\boldsymbol{\theta}_u^*) = \boldsymbol{\Sigma}_u(\boldsymbol{\theta}_u)$ if and only if $\boldsymbol{\theta}_u^* = \boldsymbol{\theta}_u$ and similar relationship between $\boldsymbol{\Sigma}_\epsilon(\boldsymbol{\theta}_\epsilon)$ and $\boldsymbol{\theta}_\epsilon$ holds. To save notation, we suppress the parameters in the parenthesis and in the following, we write $\boldsymbol{\Sigma}_u^*$ for $\boldsymbol{\Sigma}_u(\boldsymbol{\theta}_u^*)$ and similarly for other parameterized matrices.

Suppose the model is not identifiable. Then $\forall\ (\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$, there exists $(\boldsymbol{\theta}_u^*, \boldsymbol{\theta}_\epsilon^*) \neq (\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$, such that $\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon = \mathbf{Z}\boldsymbol{\Sigma}_u^*\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon^*$ or equivalently

$$\mathbf{Z}\left[\boldsymbol{\Sigma}_u^* - \boldsymbol{\Sigma}_u\right]\mathbf{Z}' = \boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*. \tag{2}$$

Below we show by contradiction that $(\boldsymbol{\theta}_u^*, \boldsymbol{\theta}_\epsilon^*) \neq (\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$ is equivalent to $\boldsymbol{\theta}_u^* \neq \boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon$. Suppose the model is not identifiable and either $\{\boldsymbol{\theta}_u^* = \boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon\}$ or $\{\boldsymbol{\theta}_u^* \neq \boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon^* = \boldsymbol{\theta}_\epsilon\}$ holds. Then by assumption, either $\{\boldsymbol{\Sigma}_u^* = \boldsymbol{\Sigma}_u, \boldsymbol{\Sigma}_\epsilon^* \neq \boldsymbol{\Sigma}_\epsilon\}$ or $\{\boldsymbol{\Sigma}_u^* \neq \boldsymbol{\Sigma}_u, \boldsymbol{\Sigma}_\epsilon^* = \boldsymbol{\Sigma}_\epsilon\}$ holds. By (2) and the full rank assumption of $\mathbf{Z}$, either $\boldsymbol{\Sigma}_u^* = \boldsymbol{\Sigma}_u$ or $\boldsymbol{\Sigma}_\epsilon^* = \boldsymbol{\Sigma}_\epsilon$ will give us $\{\boldsymbol{\Sigma}_u^* = \boldsymbol{\Sigma}_u, \boldsymbol{\Sigma}_\epsilon^* = \boldsymbol{\Sigma}_\epsilon\}$, and thus contradicts $\boldsymbol{\Sigma}_\epsilon^* \neq \boldsymbol{\Sigma}_\epsilon$ or $\boldsymbol{\Sigma}_u^* \neq \boldsymbol{\Sigma}_u$.

In practice, usually there is more than one response vector and the i-th response vector is modeled as in (1)

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\epsilon}_i,\ i = 1, \ldots, N,$$

$$\mathbf{u}_i \sim (\mathbf{0}, \boldsymbol{\Sigma}_u), \quad \boldsymbol{\epsilon}_i \sim (\mathbf{0}, \boldsymbol{\Sigma}_{\epsilon i}),$$

$$\mathbf{u}_1, \ldots, \mathbf{u}_N, \boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_N \text{ are mutually independent.} \tag{3}$$

The $\mathbf{y}_i$'s can come from an unbalanced design and the covariance matrices $\boldsymbol{\Sigma}_{\epsilon i}$'s are not necessarily equal. We call the model of all $\mathbf{y}_i$'s, (3), the *joint* model. Parameters in this model are then $(\boldsymbol{\theta}_u, \boldsymbol{\theta}_{\epsilon i},\ i = 1, \ldots, N)$. Although (3) can be expressed in the form (1) by stacking the $\mathbf{y}_i$'s into a column vector, here we make the distinction and call (1) an *individual* model. Model (1) is our target model for studying identifiability. We will investigate the identifiability relationship between an individual model and the joint model in Section 4.5, where we show the joint model is identifiable if and only if at least one individual model is identifiable.

In the following, we let $\mathbf{1}$ be a vector with each element being one and let $\mathbf{J}$ be a matrix of ones, i.e. $\mathbf{J} = \mathbf{1}\mathbf{1}'$. We let $\mathbf{I}$ be an identity matrix. Sometimes we put a subscript indicating the dimension. We let $\otimes$ denote the Kronecker product. Compound symmetry (CS) is one important covariance structure we will consider. In our study, we adopt the formulation that a matrix $\boldsymbol{\Sigma}$ has the CS structure if $\boldsymbol{\Sigma} = \sigma^2\left[(1-\rho)\mathbf{I} + \rho\mathbf{J}\right]$, for some $\sigma^2 > 0$ and $-1/(n-1) < \rho < 1$. The restrictions on σ^2 and ρ guarantee positive definiteness of $\boldsymbol{\Sigma}$ by [4]. See also Lemma A.1 in the proof of Corollary 4.1.

3. Simulations

SAS and R are widely used statistical software for data analysis and computing. The MIXED procedure of SAS and the function "lme" of R are usually used to fit an LME under normality assumption of distributions. Restricted maximum likelihood estimates (REML) are produced by default. Throughout this section, we assume all random variables are normally distributed. When fitting the model in Example 5.3, zero or extremely large standard error was observed in SAS output, and construction of confidence intervals failed using *R* [18]. When fitting a model where one of the random effects terms is totally confounded with the residual error, an extremely wide confidence interval was observed for the residual error [13].

We conducted a simulation study to examine the frequency of non-convergence, zero or extremely large standard errors, failure of confidence interval construction and unreasonably wide confidence interval bounds in non-identifiable models. In the settings we tried, which will be described in detail below, most of the times the fitting algorithms converged and about half of the times the output did not look unreasonable. We use PROC MIXED of SAS (Version 9.2) and the "lme" function of the package *nlme* in a 2.10.1 R environment on a Windows Vista platform to fit the models. We then use the R "intervals" function to

obtain the 95% confidence intervals for the standard deviation of the covariance parameters.

We simulate data from three non-identifiable models, two of which are similar to the model in [18] (revisited in Example 5.3) with a generalized design matrix associated with the random effects. Non-identifiability is justified by Corollary 4.1, (a) and (b) of Theorems 4.3 respectively, together with Theorem 4.5. In the form of (3), the response vector $\mathbf{y}_i$ is of length 2 and the matrix $\mathbf{X}_i$ is $\left(\begin{smallmatrix}1 & 0\\ 1 & 1\end{smallmatrix}\right)$ for all i. The fixed effects take value $\boldsymbol{\beta} = (1, 0.5)'$ and we have $\boldsymbol{\Sigma}_{\epsilon i} = \boldsymbol{\Sigma}_\epsilon$ for all i. In Model A, we have an MI structured $\boldsymbol{\Sigma}_u$ and a CS structured $\boldsymbol{\Sigma}_\epsilon$. In Model B, $\boldsymbol{\Sigma}_u$ is VC structured and $\boldsymbol{\Sigma}_\epsilon$ has a MI structure. In Model C, both $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_\epsilon$ are MI structured. Specifically, the models have the following form.

$$\text{Model A: } \mathbf{Z}_i = \mathbf{1}_2, \boldsymbol{\Sigma}_u = \sigma_u^2, \boldsymbol{\Sigma}_\epsilon = \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

$$\text{where } \sigma_u^2 = 1, \sigma^2 = 0.5, \rho = 0.5;$$

$$\text{Model B: } \mathbf{Z}_i = \begin{pmatrix} 1 & 1.5 \\ 1 & -1.5 \end{pmatrix}, \boldsymbol{\Sigma}_u = \begin{pmatrix} \sigma_{u1}^2 & 0 \\ 0 & \sigma_{u2}^2 \end{pmatrix}, \boldsymbol{\Sigma}_\epsilon = \sigma^2 \mathbf{I}_2,$$

$$\text{where } \sigma_{u1}^2 = 1, \sigma_{u2}^2 = 0.8, \sigma^2 = 0.5;$$

$$\text{Model C: } \mathbf{Z}_i = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \boldsymbol{\Sigma}_u = \sigma_u^2 \mathbf{I}_2, \boldsymbol{\Sigma}_\epsilon = \sigma^2 \mathbf{I}_2,$$

$$\text{where } \sigma_u^2 = 1, \sigma^2 = 0.5.$$

We consider two sample sizes $N = 500$ and $N = 1000$, and conduct 1000 simulations for each sample size and model. In the following, we show the results from the REML estimates. The estimates from the maximum likelihood method are very similar and are summarized in Tables 1-3 and Figures 7-12 of the supplementary material. Table 1 shows how many times the model fitting algorithm converges among the 1000 simulations. In the table, we see the majority converges. Table 1 also shows how many successes of constructing the confidence intervals. We see more confidence intervals are obtained for larger sample size, except for Model A. Usually intervals are obtained for more than half of the times. In particular, we get 958 intervals from Model C when $N = 1000$.

To study the scale of the standard errors of each covariance parameter, we look at the quantiles of the standard errors and summarize them in Table 2. In

TABLE 1
Number of convergence and the number of successfully constructed confidence intervals out of 1000 *simulations from REML method*

Software		N	Model A	Model B	Model C
SAS	# of Convergence	500	1000	1000	1000
		1000	1000	1000	1000
R	# of Convergence	500	1000	933	1000
		1000	1000	984	997
	# of Intervals Constructed	500	574	505	439
		1000	481	625	958

TABLE 2. *Quantiles of the standard errors for each covariance parameter from REML method*

Model	Parameter	N	0%	1%	5%	50%	60%	80%	90%	95%	99%	100%
A	$\sigma_u^2 = 1$	500	0.071	0.075	0.079	0.088	0.089	0.092	0.095	0.097	0.101	0.105
		1000	0.053	0.056	0.057	0.062	0.063	0.064	0.066	0.067	0.069	5.685×10^{4}
	$\rho = 0.5$	500	0	0	0	0	0	0	0	0	0	0
		1000	0	0	1×10^{-4}	6×10^{-4}	6×10^{-4}	7×10^{-4}	8×10^{-4}	8×10^{-4}	0.003	5.685×10^{4}
	$\sigma^2 = 0.5$	500	0.013	0.014	0.014	0.016	0.016	0.017	0.017	0.018	0.018	0.019
		1000	0.010	0.010	0.011	0.011	0.012	0.012	0.012	0.012	0.013	0.013
B	$\sigma_{u1}^2 = 1$	500	0.063	0.069	0.071	0.079	0.081	0.083	0.086	0.087	0.090	0.093
		1000	0	0	0	0	0.118	0.473	0.664	0.820	1.319	2.538
	$\sigma_{u2}^2 = 0.8$	500	0.047	0.049	0.052	0.058	0.059	0.061	0.063	0.064	0.066	0.071
		1000	0.039	0.041	0.045	0.050	0.061	0.215	0.296	0.366	0.585	1.129
	$\sigma^2 = 0.5$	500	0	0	0	0	0	0	0	0	0	0
		1000	0	0.024	0.102	0.117	0.208	0.947	1.328	1.636	2.632	5.073
C	$\sigma_u^2 = 1$	500	0.050	0.051	0.052	0.056	0.057	0.058	0.059	0.060	0.062	0.064
		1000	0	0	0	0	0.154	0.468	0.653	0.874	1.307	2.002
	$\sigma^2 = 0.5$	500	0	0	0	0	0	0	0	0	0	0
		1000	0.004	0.023	0.075	0.083	0.298	0.933	1.305	1.748	2.615	4.003

the table, we see when $N = 500$, parameters ρ in Model A, σ^2 in Model B and σ^2 in Model C have identically zero standard errors. However, for bigger sample size $N = 1000$, there is only a tiny proportion of zero standard errors for them. Though about 50% of the standard errors of σ^2_{u1} and σ^2_u are zero in Model B and C respectively, the rest of the other 50% are not of a large scale with the maximum being 2.538 and 2.002. We observe extremely large standard errors for σ^2_u and ρ in Model A when $N = 1000$, but the proportion of them is no more than 1%. Histograms for each estimated parameters are provided in Figures 1, 3 and 5 of the supplementary material. In the plots, we do not observe large scale estimates but sometimes zero estimates where the standard errors are also zero.

To check the length of the 95% confidence intervals on the covariance parameters, we summarize the quantiles of the lengths in Table 3. In the table, we observe smaller lengths for bigger sample size $N = 1000$. In Model A, when $N = 500$, 90% of the intervals are shorter than 4 with the widest being 3.989 for σ. When $N = 1000$, the proportion increases to 95% by the same cutoff 4. In Model B, we don't observe very wide intervals for σ_{u1} and σ_{u2} when $N = 1000$ and 99% of the intervals are not wide with the widest being 3.845 when $N = 500$. For σ, 80% of the intervals are shorter than 3.18 when $N = 500$ and 95% of them are shorter than 1.917 when $N = 1000$. In Model C, for σ_u, 95% of the intervals are shorter than 1.435 when $N = 500$ while all the intervals are shorter than 2.757 when $N = 1000$. For σ, 60% of the intervals are shorter than 3.405 when $N = 500$ and 95% are shorter than 1.671 when $N = 1000$. Histograms for each estimated parameters are provided in Figures 2, 4 and 6 of the supplementary material.

In summary, for the non-identifiable models considered in the simulation study, the fact that they were non-identifiable could often not be easily detected, and it would be useful to have another way to check identifiability. We will provide an easy to use analytic method in the next Section.

4. Main results

In this section, we present conditions of identifiability and study commonly used covariance structures. In Sections 4.1-4.4, our study focuses on the individual model (1). In Section 4.5, we investigate the relationship of identifiability between the individual model (1) and the joint model (3), and generalize a known result in [2].

The obtained results in Sections 4.1-4.4 are organized as follows. Theorem 4.1 gives a necessary and sufficient condition for the existence of $(\boldsymbol{\theta}^*_u, \boldsymbol{\theta}^*_\epsilon) \neq (\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$, satisfying (2). The condition relies mainly on the design matrix $\mathbf{Z}$ via $\mathbf{H}_\mathbf{Z} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$. A sufficient condition of identifiability is then provided in Theorem 4.2. Necessary and sufficient conditions of non-identifiability under a UN $\boldsymbol{\Sigma}_u$ and a CS or VC $\boldsymbol{\Sigma}_\epsilon$ are derived from Theorem 4.1 and presented in Section 4.2. In Section 4.3, identifiability conditions under an MK or MA(1) $\boldsymbol{\Sigma}_\epsilon$ are derived from Theorem 4.2. In Section 4.4, we derive conditions of non-identifiability under a VC or MI $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_\epsilon$ from Theorem 4.1, where $\mathbf{Z}$ has a more general form than in Example 5.3.

TABLE 3. *Quantiles of the length of the 95% confidence intervals on the covariance parameters from REML method*

Model	Parameter	N	0%	1%	5%	50%	60%	80%	90%	95%	99%	100%
A	$\sigma_u = 1$	500	0.353	0.380	0.441	0.761	0.871	1.331	2.049	3.506	91.004	8.113×10^{11}
		1000	0.182	0.205	0.233	0.419	0.481	0.696	0.979	1.544	3.153	18.322
	$\rho = 0.5$	500	0.706	0.748	0.868	1.325	1.472	1.764	1.940	1.994	2	2
		1000	0.515	0.606	0.667	1.107	1.242	1.565	1.810	1.969	1.999	2
	$\sigma = \sqrt{0.5}$	500	0.532	0.555	0.657	1.154	1.378	2.220	3.989	10.165	1.660×10^{4}	2.215×10^{21}
		1000	0.289	0.349	0.383	0.698	0.823	1.263	1.875	3.935	14.072	5.447×10^{2}
B	$\sigma_{u1} = 1$	500	0.180	0.193	0.220	0.367	0.427	0.659	0.955	1.332	3.845	4.000×10^{2}
		1000	0.090	0.099	0.110	0.172	0.188	0.251	0.344	0.486	1.154	2.018
	$\sigma_{u2} = \sqrt{0.8}$	500	0.136	0.155	0.163	0.219	0.248	0.346	0.493	0.674	1.879	45.131
		1000	0.079	0.087	0.096	0.120	0.126	0.153	0.204	0.262	0.547	0.897
	$\sigma = \sqrt{0.5}$	500	0.591	0.708	0.779	1.354	1.676	3.180	6.529	17.589	1.301×10^{4}	4.684×10^{14}
		1000	0.335	0.349	0.386	0.588	0.642	0.881	1.340	1.917	13.267	46.253
C	$\sigma_u = 1$	500	0.245	0.265	0.291	0.488	0.549	0.802	1.089	1.435	6.429	1.385×10^{2}
		1000	0.100	0.107	0.114	0.147	0.161	0.203	0.262	0.346	0.712	2.757
	$\sigma = \sqrt{0.5}$	500	1.042	1.159	1.322	2.761	3.405	7.898	19.075	48.502	2.871×10^{5}	5.077×10^{14}
		1000	0.410	0.436	0.471	0.602	0.660	0.850	1.117	1.671	5.715	1.631×10^{3}

4.1. General conditions

Theorem 4.1. *$\forall(\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$, $\exists\ (\boldsymbol{\theta}_u^*, \boldsymbol{\theta}_\epsilon^*) \neq (\boldsymbol{\theta}_u, \boldsymbol{\theta}_\epsilon)$ such that (2) holds if and only if $\forall \boldsymbol{\theta}_\epsilon$, $\exists\ \boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon$, such that $\mathbf{Z}'\boldsymbol{\Sigma}_\epsilon\mathbf{Z} \neq \mathbf{Z}'\boldsymbol{\Sigma}_\epsilon^*\mathbf{Z}$,*

$$\mathbf{H}_\mathbf{Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right] = \boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*, \tag{4}$$

and

$$\boldsymbol{\Sigma}_u^* = \boldsymbol{\Sigma}_u + (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}. \tag{5}$$

We note that $\boldsymbol{\theta}_u^* \neq \boldsymbol{\theta}_u$ is implied by the first condition and (5). Theorem 4.1 serves as a necessary condition for model non-identifiability. When the model is not identifiable, (5) also gives the form of $\boldsymbol{\Sigma}_u^*$. Usually parameters in $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_\epsilon$ are unknown, and the existence of $\boldsymbol{\theta}_\epsilon^*$ remains uncertain. However, (4) or (5) can be verifiable irrespective of the unknowns. If any of the conditions fails, the model is then identifiable. In Example 5.2, we see (4) fails where $\boldsymbol{\Sigma}_\epsilon$ has a block diagonal form. If one finds a $\boldsymbol{\theta}_\epsilon^*$ satisfying the conditions, the model is then not identifiable if $\boldsymbol{\theta}_\epsilon^* \in \boldsymbol{\Theta}_\epsilon$ and $\boldsymbol{\theta}_u^* \in \boldsymbol{\Theta}_u$. The results in Sections 4.2 and 4.4 are then derived from this perspective where we find a $\boldsymbol{\theta}_\epsilon^*$ deviating from $\boldsymbol{\theta}_\epsilon$ by one or more elements. Once the structures are specified, non-identifiability only relies on whether certain properties of $\mathbf{Z}$ are satisfied or not.

Suppose $\forall\boldsymbol{\theta}_\epsilon$, $\exists\ \boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon$ satisfying (4). Since the rank of $\mathbf{H}_\mathbf{Z}$ is q, the rank of $\mathbf{H}_\mathbf{Z}(\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*)$ is at most q. In order for (4) to hold, the rank of $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*$ has to be less than or equal to q. If for any $\boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon$ which are both in $\boldsymbol{\Theta}_\epsilon$, we have $\operatorname{rank}(\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*) > q$, (4) then fails.

Theorem 4.2. *Suppose $n > q$. Model (1) is identifiable if $\operatorname{rank}(\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*) > q$ for any $\boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon$ which are both in $\boldsymbol{\Theta}_\epsilon$.*

4.2. UN+CS/VC

We derive if and only if conditions of non-identifiability under a UN $\boldsymbol{\Sigma}_u$ and a CS or VC $\boldsymbol{\Sigma}_\epsilon$.

Corollary 4.1. *Suppose $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_\epsilon$ have the UN and the CS structures respectively. Suppose the matrix $\mathbf{Z}$ satisfies $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$ and $\operatorname{rank}(\mathbf{Z}) = q$ with $1 \leq q < n-1$. Model (1) is non-identifiable if and only if $\mathbf{H}_\mathbf{Z}\mathbf{1} = \mathbf{1}$.*

The condition $\mathbf{H}_\mathbf{Z}\mathbf{1} = \mathbf{1}$ amounts to checking if the sum of the elements of each row of $\mathbf{H}_\mathbf{Z}$ is equal to one. For the case $q = 1$, $\boldsymbol{\Sigma}_u$ is a scalar and $\mathbf{Z}$ is a column vector $(z_1, \ldots, z_n)'$ with

$$\mathbf{H}_\mathbf{Z} = \begin{pmatrix} \frac{z_1^2}{s_z^2} & \frac{z_1 z_2}{s_z^2} & \cdots & \frac{z_1 z_n}{s_z^2} \\ \vdots & & & \vdots \\ \frac{z_n z_1}{s_z^2} & \frac{z_n z_2}{s_z^2} & \cdots & \frac{z_n^2}{s_z^2} \end{pmatrix}, \quad s_z^2 = \sum z_i^2.$$

Thus, for $q = 1$, the model is non-identifiable if and only if $\mathbf{Z}$ is a non-zero constant vector.

Suppose that $\mathbf{\Sigma}_\epsilon$ have a VC structure as

$$\mathbf{\Sigma}_\epsilon = \operatorname{diag}\{\sigma_1^2\mathbf{I}_{n_1}, \sigma_2^2\mathbf{I}_{n_2}, \sigma_k^2\mathbf{I}_{n_k}\},\ \sum_{j=1}^{k} n_j = n,\ \sigma_j^2 > 0,\ j = 1, \ldots, k,$$

$$\sigma_i^2 \neq \sigma_j^2, \quad \text{if } i \neq j. \tag{6}$$

Corresponding to the dimension n_j of each component σ_j^2, we first partition the matrix $\mathbf{Z}$ into sub-matrices $\mathbf{Z}_j$ with dimensions $n_j \times q$, $j = 1, \ldots, k$, i.e. $\mathbf{Z}' = (\mathbf{Z}_1', \ldots, \mathbf{Z}_k')$. Similarly, we partition diagonal elements of $\mathbf{H}_\mathbf{Z}$ into vectors, $\mathbf{h}_j$'s, where the length of $\mathbf{h}_j$ is n_j, $1 \leq j \leq k$. That is, $\operatorname{diag}(\mathbf{H}_\mathbf{Z}) = (\mathbf{h}_1', \ldots, \mathbf{h}_k')'$.

Corollary 4.2. *Suppose $\mathbf{\Sigma}_u$ and $\mathbf{\Sigma}_\epsilon$ are UN and VC structured respectively. Model (1) is not identifiable if and only if for at least one j, $\mathbf{Z}_j'\mathbf{Z}_j \neq \mathbf{0}$ and $\mathbf{h}_j = \mathbf{1}$, $1 \leq j \leq k$.*

4.3. MK/MA(1)

We will see the model is often identifiable as the conditions imposed on the dimension of $\mathbf{Z}$ are very mild, irrespective to the $\mathbf{\Sigma}_u$ structure. We consider the MK $\mathbf{\Sigma}_\epsilon$ as $\mathbf{\Sigma}_\epsilon = \sigma^2\mathbf{R}$, where $\mathbf{R}$ is a known covariance matrix. Clearly for any $\sigma^{*2} \neq \sigma^2$, $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{R}$ is invertible, and so is of rank n. By Theorem 4.2, we get the following result.

Corollary 4.3. *Suppose $n > q$. Model (1) is identifiable if $\mathbf{\Sigma}_\epsilon = \sigma^2\mathbf{R}$, where $\sigma^2 > 0$ and $\mathbf{R}$ is a known covariance matrix.*

Let $\mathbf{T}$ be an $n \times n$ Toeplitz matrix with ones on the two parallel subdiagonals and zeroes elsewhere. An MA(1) structured $\mathbf{\Sigma}_\epsilon$ has the form $\sigma^2(\mathbf{I} + \rho\mathbf{T})$, $\sigma^2 > 0$, $|\rho| < 1/2$. We have the following result.

Corollary 4.4. *Suppose $n - 1 > q$. Model (1) is identifiable under an MA(1) structured $\mathbf{\Sigma}_\epsilon$.*

4.4. Conditions under a generalized $\mathbf{Z}$

We consider a more general design matrix

$$\mathbf{Z} = \begin{pmatrix} 1 & v \\ 1 & z \end{pmatrix},\ z \neq v, \tag{7}$$

than $\left(\begin{smallmatrix} 1 & 3 \\ 1 & 6 \end{smallmatrix}\right)$ in Example 5.3. We derive conditions under which the model is not identifiable where $\mathbf{\Sigma}_u$ or $\mathbf{\Sigma}_\epsilon$ is MI or VC structured. We note that under these structures $\mathbf{\Sigma}_y$ is not over-parametrized as the total number of parameters does not exceed 3.

Theorem 4.3. *Given $\mathbf{Z}$ in (7), the model is not identifiable in each of the following cases: (a) $z = -v$, $\mathbf{\Sigma}_u$ has a VC structure and $\mathbf{\Sigma}_\epsilon$ has an MI structure; (b) $z = -v$, $z^2 = 1$, both $\mathbf{\Sigma}_u$ and $\mathbf{\Sigma}_\epsilon$ have MI structures; (c) $zv = -1$, $\mathbf{\Sigma}_u$ is MI structured and $\mathbf{\Sigma}_\epsilon$ has a VC structure.*

4.5. Conditions of the joint model

Statistical inference is normally based on the joint model (3) of all the individuals. Intuitively, if we are able to identify $\mathbf{\Sigma}_u$ from one individual model, then we can identify all of the $\mathbf{\Sigma}_{\epsilon i}$'s. Theorem 4.4 summaries this observation for an individual model and will be used in the proof of Theorem 4.5.

Theorem 4.4. *The model (1) is identifiable under any of the following conditions: (a)* $\mathbf{\Sigma}_u$ *is known; (b)* $\mathbf{\Sigma}_\epsilon$ *is known; (c)* $\mathbf{Z}\mathbf{\Sigma}_u\mathbf{Z}'{\mathbf{\Sigma}_\epsilon}^{-1} = \mathbf{K}$*, where* $\mathbf{K}$ *is known and* $\mathbf{K}+\mathbf{I}$ *is of full column rank.*

The conditions are analogous to those in measurement error models where measurement variance or ratio of measurement variances is assumed known [3] (p. 14, 30, 103).

Theorem 4.5. *The joint model (3) is identifiable if and only if at least one individual model is identifiable.*

Theorem 4.5 reduces identifiability verification of a joint model to the verification of individual models. For instance, if an i-th individual model has a $n_i \times q$ $\mathbf{Z}_i$ of full rank and $\mathbf{\Sigma}_{\epsilon i} = \sigma^2\mathbf{I}_{n_i}$, where $n_i > q$, then this individual model is identifiable by Corollary 4.3 and so is the joint model. Note that the other individual models can still have their $\mathbf{Z}_j$'s not of full rank.

[2] studies identifiability of the joint model assuming $\mathbf{\Sigma}_{\epsilon i} = \sigma^2\mathbf{I}_{n_i}$ and $\mathbf{\Sigma}_u = \sigma^2\mathbf{E}$, where $\mathbf{E}$ is an unknown covariance matrix. The author shows that the joint model is identifiable if at least one matrix $\mathbf{Z}_i$ is of full rank and $\sum_{i=1}^{N}(n_i - q) > 0$. This setup in which covariance matrices $\mathbf{\Sigma}_{\epsilon i}$ and $\mathbf{\Sigma}_u$ share a common parameter σ^2 is different from ours. We observe that the condition $\sum_{i=1}^{N}(n_i - q) > 0$ implies $n_i > q$ for at least one i while the reverse is not necessarily true. Our result can also be applied to study $\mathbf{\Sigma}_{\epsilon i}$'s of other structures.

5. Examples

We verify non-identifiability or identifiability of models in literature using the results in Section 4. For simplicity, we only give brief descriptions, and omit the fixed effects part in the models and only write the random effects and the error part, indicated by $\overset{r}{=}$. For details of the studies, please refer to the corresponding literature.

Example 5.1.

[9] analyzed data from an experiment that investigated effects of several supplemental sources of dietary Mg (Magnesium) on urinary Mg excretion in lambs. Lambs were assigned to each treatment and urinary Mg was measured for each lamb on consecutive days. Let y_{ijk} denote the response at time k from lamb j in treatment group i. Let u_{ij} denote the random effect of lamb j in treatment group i and let ϵ_{ijk} denote the error. The model used for the repeated measures experiment is

$$y_{ijk} \overset{r}{=} u_{ij} + \epsilon_{ijk},\ k = 1, \ldots, t,$$

$$u_{ij} \sim \mathrm{N}(0, \sigma_u^2),\ \boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \ldots, \epsilon_{ijt})' \sim \mathrm{N}(0, \boldsymbol{\Sigma}_\epsilon),$$

$$\ldots, u_{ij}, \ldots, \boldsymbol{\epsilon}_{ij}, \ldots, \text{ are mutually independent.}$$

Let $\mathbf{y}_{ij} = (y_{ij1}, \ldots, y_{ijt})'$. We write the model as

$$\mathbf{y}_{ij} \overset{r}{=} \mathbf{1}_t u_{ij} + \boldsymbol{\epsilon}_{ij}.$$

UN and CS $\boldsymbol{\Sigma}_\epsilon$ were considered in the model. We notice that a UN $\boldsymbol{\Sigma}_\epsilon$ with an additional σ_u^2 over-parameterizes the covariance matrix of $\mathbf{y}_{ij}$. Although there is no over-parameterization under the CS structure, non-identifiability of an individual model follows from Corollary 4.1 and its following observation. By Theorem 4.5, the joint model of all $\mathbf{y}_{ij}$'s is not identifiable either.

Example 5.2.

[14] and [19] studied the effect of different lead treatments on the growths of American Kestrel nestlings. The American Kestrel nestlings, in each of ten nests, are orally dosed with a treatment and the birds of each nest were weighed every day. The experiment followed a randomized block design with nests as the blocking factor. Let y_{ijk} denote the log(body weight) of the jth bird ($j = 1, 2, 3$) in the ith nest ($i = 1, \ldots, 10$) on the kth day ($k = 1, \ldots, 5$). The model is

$$y_{ijk} \overset{r}{=} u_{ik} + \epsilon_{ijk},\ k = 1, \ldots, 5,$$

$$\mathbf{u}_i = (u_{i1}, \ldots, u_{i5})' \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_u), \boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \ldots, \epsilon_{ij5})' \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon),$$

$$\ldots, \mathbf{u}_i, \ldots, \boldsymbol{\epsilon}_{ij}, \ldots \text{ are mutually independent.}$$

This model is the same as the model in Example 5.1 except that the random effect u_{ik} relies on the index k of the repeated measures instead of on the index j of the individuals.

Let $\mathbf{y}_{ij} = (y_{ij1}, \ldots, y_{ij5})'$ and let $\mathbf{y}_i = (\mathbf{y}_{i1}', \ldots, \mathbf{y}_{i3}')'$. We write

$$\mathbf{y}_i \overset{r}{=} (\mathbf{1}_3 \otimes \mathbf{I}_5)\, \mathbf{u}_i + \boldsymbol{\epsilon}_i,$$

$$\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}_{i1}', \ldots, \boldsymbol{\epsilon}_{i3}')' \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_3 \otimes \boldsymbol{\Sigma}_\epsilon).$$

For this model, we have $\mathbf{Z} = \mathbf{1}_3 \otimes \mathbf{I}_5$, $\mathbf{Z}'\mathbf{Z} = 3\mathbf{I}_5$, and $\mathbf{H}_\mathbf{Z} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}' = (\mathbf{J}_3/3) \otimes \mathbf{I}_5$. The left hand side of (4) in Theorem 4.1 is equal to $\mathbf{H}_\mathbf{Z}[\mathbf{I}_3 \otimes (\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*)] = (\mathbf{J}_3/3) \otimes (\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*)$. In order for this to equal the right hand side $\mathbf{I}_3 \otimes (\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*)$, we must have $\boldsymbol{\theta}_\epsilon = \boldsymbol{\theta}_\epsilon^*$. The condition (4) is not satisfied and thus the model of a $\mathbf{y}_i$ is identifiable. By Theorem 4.5, the joint model of all $\mathbf{y}_i$'s is also identifiable.

Example 5.3.

[18] analyzed a Dental Veneer data set where the response, gingival crevicular fluid (GCF), was measured at two post-treatment time points ($t_1 = 3$ and $t_2 = 6$ months) for each tooth, with teeth nested within patients. Let y_{ijk} denote an

individual GCF response at visit k ($k = 1, 2$, corresponding to months 3 and 6) on tooth j nested within patient i. The model is

$$y_{ijk} \overset{r}{=} (1 \;\; t_k)\,\mathbf{v}_i + u_{ij} + \epsilon_{ijk},\;\; k = 1, 2,$$

$$\mathbf{v}_i = (v_{i1}, v_{i2})' \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_v),\;\; u_{ij} \sim \mathrm{N}(0, \sigma_u^2),\;\; \boldsymbol{\epsilon}_{ij} = (\epsilon_{ij1}, \epsilon_{ij2})' \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_\epsilon),$$

$$\ldots, \mathbf{v}_i, \ldots, u_{ij}, \ldots, \boldsymbol{\epsilon}_{ij}, \ldots, \text{ are mutually independent.}$$

Let $\mathbf{y}_{ij} = (y_{ij1}, y_{ij2})$ and $\mathbf{Z}_v = \left(\begin{smallmatrix} 1 & 3 \\ 1 & 6 \end{smallmatrix}\right)$. We write

$$\mathbf{y}_{ij} \overset{r}{=} \mathbf{Z}_v \mathbf{v}_i + \mathbf{1}_2 u_{ij} + \boldsymbol{\epsilon}_{ij},$$

and get $\boldsymbol{\Sigma}_{y_{ij}} = \mathbf{Z}_v \boldsymbol{\Sigma}_v \mathbf{Z}_v' + \boldsymbol{\Sigma}_{ij}$, where $\boldsymbol{\Sigma}_{ij} = \sigma_u^2 \mathbf{J}_2 + \boldsymbol{\Sigma}_\epsilon$.

Let $\mathbf{y}_i = (\mathbf{y}_{i1}, \ldots, \mathbf{y}_{in_i})$. We further write

$$\mathbf{y}_i \overset{r}{=} (\mathbf{1}_{n_i} \otimes \mathbf{Z}_v)\,\mathbf{v}_i + (\mathbf{I}_{n_i} \otimes \mathbf{1}_2)\,\mathbf{u}_i + \boldsymbol{\epsilon}_i,$$

$$\mathbf{u}_i = (u_{i1}, \ldots, u_{in_i})' \sim \mathrm{N}(\mathbf{0}, \sigma_u^2 \mathbf{I}_{n_i}),\;\; \boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}_{i1}', \ldots, \boldsymbol{\epsilon}_{in_i}')' \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_\epsilon).$$

UN and CS $\boldsymbol{\Sigma}_\epsilon$ were considered and we show non-identifiability of the models. We consider the random part of $\mathbf{y}_i$ leaving out the term involving $\mathbf{v}_i$, and get $\mathrm{Cov}((\mathbf{I}_{n_i} \otimes \mathbf{1}_2)\,\mathbf{u}_i + \boldsymbol{\epsilon}_i) = \mathbf{I}_{n_i} \otimes \sigma_u^2(\mathbf{1}_2\mathbf{1}_2') + \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_\epsilon = \mathbf{I}_{n_i} \otimes \boldsymbol{\Sigma}_{ij}$. Similarly as in Example 5.1, $\boldsymbol{\Sigma}_{ij}$ is not identifiable under a UN or CS $\boldsymbol{\Sigma}_\epsilon$, and so is the covariance matrix of $\mathbf{y}_i$.

6. Summary and discussion

We derive conditions of identifiability for the covariance parameters in an LME and study commonly used covariance structures. Theorem 4.1 serves as a necessary condition of non-identifiability. The conditions can be verifiable irrespective of the structures specified. If any of the conditions fails, the model is identifiable. If the conditions are satisfied, identifiability depends on the structures used and if certain properties of the design matrix associated with the random effects, $\mathbf{Z}$, are satisfied or not. Theorem 4.2 presents a sufficient condition of identifiability. To check identifiability of the joint model, it reduces to checking if at least one individual model is identifiable.

Our study is based on model (1) with two random components, $\mathbf{u}$ and $\boldsymbol{\epsilon}$. In some models, there may be three random components as in Example 5.3 or more. The results may still be applicable to study identifiability of these models. For instance, suppose there are two random effects and using similar notation, we have

$$\mathbf{y} \overset{r}{=} \mathbf{Z}_1 \mathbf{u}_1 + \mathbf{Z}_2 \mathbf{u}_2 + \boldsymbol{\epsilon},$$

$$\mathbf{u}_1 \sim (\mathbf{0}, \boldsymbol{\Sigma}_{u1}),\;\; \mathbf{u}_2 \sim (\mathbf{0}, \boldsymbol{\Sigma}_{u2}).$$

If $\mathbf{u}_1$ and $\mathbf{u}_2$ are independent, we can group one of them with the error, and study identifiability of the grouped term alone and identifiability with the grouped term and the remaining random effects included.

On the other hand, the model can be written in the form of (1) with

$$\mathbf{Z} = (\mathbf{Z}_1\ \mathbf{Z}_2)\,,\ \mathbf{u} = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix},\ \boldsymbol{\Sigma}_u = \begin{pmatrix} \boldsymbol{\Sigma}_{u1} & \boldsymbol{\Sigma}_{u1,u2} \\ \boldsymbol{\Sigma}_{u2,u1} & \boldsymbol{\Sigma}_{u2} \end{pmatrix}.$$

If $\boldsymbol{\Sigma}_{u1,u2} \neq \mathbf{0}$, $\boldsymbol{\Sigma}_u$ is UN structured and the conditions under a CS or VC $\boldsymbol{\Sigma}_\epsilon$ can be verifiable. If $\boldsymbol{\Sigma}_{u1,u2} = \mathbf{0}$, $\boldsymbol{\Sigma}_u$ then has a block diagonal structure. The conditions under an MK or MA(1) $\boldsymbol{\Sigma}_\epsilon$ can still be checked as only the dimension of $\mathbf{Z}$ is involved. Other conditions may be derived from Theorem 4.1 for specific $\mathbf{Z}$ and structures.

Supplementary Material

Supplementary materials for identifiability of linear mixed effects models

(doi: 10.1214/13-EJS770SUPP; .pdf). Tables and figures referenced in Section 3 are provided in the supplementary material.

Acknowledgements

The author would like to thank the referee and associate editor for very helpful comments that have improved the paper.

Appendix

Theorem 4.1

Proof. Suppose the model is not identifiable. We premultiply (2) by $\mathbf{Z}'$, postmultiply it by $\mathbf{Z}$ and then pre- and postmultiply by $(\mathbf{Z}'\mathbf{Z})^{-1}$ to get

$$\boldsymbol{\Sigma}_u^* - \boldsymbol{\Sigma}_u = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}, \tag{8}$$

which gives (5). Given $\boldsymbol{\Sigma}_u^* \neq \boldsymbol{\Sigma}_u$, we have $\mathbf{Z}'\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{Z} \neq \mathbf{0}$. To derive (4), premultiply (8) by $\mathbf{Z}$, postmultiply (8) by $\mathbf{Z}'$ to get

$$\mathbf{Z}(\boldsymbol{\Sigma}_u^* - \boldsymbol{\Sigma}_u)\mathbf{Z}' = \mathbf{H}_\mathbf{Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{H}_\mathbf{Z} \tag{9}$$

which, by (2), is the same as

$$\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^* = \mathbf{H}_\mathbf{Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{H}_\mathbf{Z}. \tag{10}$$

Premultiplying (10) by the idempotent matrix $\mathbf{H}_\mathbf{Z}$ gives

$$\mathbf{H}_\mathbf{Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right] = \mathbf{H}_\mathbf{Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{H}_\mathbf{Z}.$$

Substituting (10) into the right side of the above yields (4).

To prove the converse, we want to show that (4) and (5) lead to (2). It is clear from (5) that (9) holds. The conditions $\mathbf{\Sigma}_\epsilon \neq \mathbf{\Sigma}_\epsilon^*$ and $\mathbf{Z}'\mathbf{\Sigma}_\epsilon\mathbf{Z} \neq \mathbf{Z}'\mathbf{\Sigma}_\epsilon^*\mathbf{Z}$ ensure $\mathbf{\Sigma}_u \neq \mathbf{\Sigma}_u^*$. It remains to show that (10) holds since substituting (10) into the right side of (9) yields (2). To show (10), from (4) and the symmetry of $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*$, we see that

$$\mathbf{H_Z}[\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*] = [\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*]\mathbf{H_Z}.$$

Premultiplying the above identity by the idempotent matrix $\mathbf{H_Z}$ gives $\mathbf{H_Z}[\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*] = \mathbf{H_Z}[\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*]\mathbf{H_Z}$. Substituting (4) for the left side of the equation gives (10). □

Corollary 4.1

Proof. To prove the corollary, we use the following result in [4].

Lemma A.1. *Given two scalars a and b, the characteristic equation of the matrix $\mathbf{C} = (a-b)\mathbf{I} + b\mathbf{J}$ in λ is*

$$(a + (n-1)b - \lambda)\,(a - b - \lambda)^{n-1},$$

and hence $n-1$ characteristic roots are equal to $a-b$ and one root is equal to $a + (n-1)b$.

Suppose $\mathbf{H_Z}\mathbf{1} = \mathbf{1}$. Given an arbitrary $\mathbf{\Sigma}_\epsilon$, let $s > 1$, $\sigma^{*2} = s\sigma^2$ and $\rho^* = (\rho - 1)/s + 1$. We get $-1/(n-1) < \rho^* < 1$. Define $\mathbf{\Sigma}_\epsilon^* = \sigma^{*2}\,[(1-\rho^*)\mathbf{I} + \rho^*\mathbf{J}]$. Then $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{J}$ and, since $\mathbf{H_Z}\mathbf{1} = \mathbf{1}$, it is clear that (4) is satisfied. By assumption $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$, it holds $\mathbf{Z}'(\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*)\mathbf{Z} \neq \mathbf{0}$. We now show that, for any $\mathbf{\Sigma}_u$ of UN structure, there exists $s^* > 1$ so that $\mathbf{\Sigma}_u^*$ defined as in (5) is also UN structured whenever $1 < s < s^*$. Plugging $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{J}$ into (5) yields

$$\mathbf{\Sigma}_u^* = \mathbf{\Sigma}_u + \sigma^2(1-s)(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{J}\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}. \tag{11}$$

By assumption $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$ and $\mathbf{Z}$ is of full column rank, the matrix $(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{J}\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}$ is non-negative definite and of rank one since $\mathbf{J} = \mathbf{1}\mathbf{1}'$. Let λ be its non-zero and thus the largest eigenvalue of $(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{J}\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}$. Let λ_m be the smallest eigenvalue of the matrix $\mathbf{\Sigma}_u$, and let $s^* = \lambda_m/(\lambda\sigma^2) + 1$.

$$\begin{aligned}\mathbf{\Sigma}_u^* &= \mathbf{\Sigma}_u + \sigma^2(1-s)(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{J}\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1} \\ &\geq \lambda_m\mathbf{I} + \sigma^2(1-s)\lambda\mathbf{I} \\ &> \mathbf{0},\end{aligned}$$

whenever $1 < s < s^*$.

Now suppose that the model is not identifiable. Then, by Theorem 4.1, $\forall\ \mathbf{\Sigma}_\epsilon$, there exists $\boldsymbol{\theta}_\epsilon^* \neq \boldsymbol{\theta}_\epsilon$ satisfying (4) and, since the rank of $\mathbf{H_Z}$ is q, the rank of $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*$ is at most q. We have

$$\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = \left[(\sigma^2 - \sigma^{*2}) - (\sigma^2\rho - \sigma^{*2}\rho^*)\right]\mathbf{I} + (\sigma^2\rho - \sigma^{*2}\rho^*)\mathbf{J}.$$

By Lemma A.1, the eigenvalues of $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*$ are $(\sigma^2 - \sigma^{*2}) - (\sigma^2\rho - \sigma^{*2}\rho^*)$, of multiplicity $n-1$ and $(\sigma^2 - \sigma^{*2}) + (n-1)(\sigma^2\rho - \sigma^{*2}\rho^*)$, of multiplicity 1. Since $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*$ is not a zero matrix, all of the eigenvalues cannot be equal to 0: we must either have no eigenvalues equal to 0, one eigenvalue equal to 0, or $n-1$ eigenvalues equal to 0. In order to have $\text{rank}(\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*) \leq q$, the $n-1$ multiple eigenvalues have to be zero since $1 \leq q < n-1$ by assumption. That is, $\sigma^2 - \sigma^{*2} = \sigma^2\rho - \sigma^{*2}\rho^*$ and so $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{J}$. Plugging this into (4) yields $\mathbf{H_Z J} = \mathbf{J}$ which further gives $\mathbf{H_Z 1} = \mathbf{1}$. □

Corollary 4.2

Proof. We first note a fact about the matrix $\mathbf{H_Z}$. Since $\mathbf{H_Z}$ is symmetric and idempotent,

$$\mathbf{H_Z}[k,k] = \sum_l (\mathbf{H_Z}[k,l])^2 = (\mathbf{H_Z}[k,k])^2 + \sum_{l \neq k} (\mathbf{H_Z}[k,l])^2 .$$

Thus, if $\mathbf{H_Z}[k,k] = 1$, then $\mathbf{H_Z}[k,l] = \mathbf{H_Z}[l,k] = 0$ for all $l \neq k$. To prove the corollary, without loss of generality, we consider the case $j = 1$.

Suppose $\mathbf{h}_1 = \mathbf{1}$. That is, $\mathbf{H_Z}[i,i] = 1$ for all $1 \leq i \leq n_1$. Then by the observation above, $\mathbf{H_Z}[i,j] = \mathbf{H_Z}[j,i] = 0$ for all $j \neq i$. Given an arbitrary $\mathbf{\Sigma}_\epsilon$ as in (6), we choose σ_1^{*2} where $0 < \sigma_1^{*2} < \sigma_1^2$, and define $\mathbf{\Sigma}_\epsilon^* = \text{diag}\{\sigma_1^{*2}\mathbf{I}_{n_1}, \sigma_2^2\mathbf{I}_{n_2}, \ldots, \sigma_k^2\mathbf{I}_{n_k}\}$. Then $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = \text{diag}\{(\sigma_1^2 - \sigma_1^{*2})\mathbf{I}_{n_1}, 0, \ldots, 0\}$. As such, (4) is satisfied. Since $(\sigma_1^2 - \sigma_1^{*2}) > 0$ and $\mathbf{Z}_1'\mathbf{Z}_1 \neq \mathbf{0}$, $\mathbf{Z}'(\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*)\mathbf{Z} = (\sigma_1^2 - \sigma_1^{*2})\mathbf{Z}_1'\mathbf{Z}_1$ is semi-positive definite. Thus, $\mathbf{\Sigma}_u^*$ defined in (5) also has a UN structure. Therefore, the model is not identifiable by Theorem 4.1.

Suppose that the model is not identifiable. Then by Theorem 4.1, $\forall\ \mathbf{\Sigma}_\epsilon$, there exists $\mathbf{\Sigma}_\epsilon^* \neq \mathbf{\Sigma}_\epsilon$ satisfying (4). That is, at least one of the $(\sigma_j^2 - \sigma_j^{*2})\mathbf{I}_{n_j}$ is not zero. Without loss of generality, we assume that $\sigma_1^2 \neq \sigma_1^{*2}$. By (4), the 1st to n_1th diagonal elements of $\mathbf{H_Z}$ must be one. That is, $\mathbf{h}_1 = \mathbf{1}$. Since $\mathbf{\Sigma}_u^* \neq \mathbf{\Sigma}_u$ which is implied by model non-identifiability, we have from (5) $\mathbf{Z}'(\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*)\mathbf{Z} = \mathbf{Z}'\text{diag}\{(\sigma_1^2 - \sigma_1^{*2})\mathbf{I}_{n_1}, 0, \ldots, 0\}\mathbf{Z} \neq \mathbf{0}$. That is, $\mathbf{Z}_1'\mathbf{Z}_1 \neq \mathbf{0}$. □

Corollary 4.4

Proof. To prove the corollary, we use the following lemma which is a result in [4]

Lemma A.2. *Let* $\mathbf{T}$ *be an* $n \times n$ *Toeplitz matrix with ones on the two parallel subdiagonals and zeroes elsewhere. Given two scalars* a_0 *and* a_1*, the eigenvalues of the* $n \times n$ *matrix* $\mathbf{C} = a_0\mathbf{I} + a_1\mathbf{T}$ *are*

$$\lambda_i = a_0 + 2|a_1| \cos\frac{i\pi}{n+1},\ i = 1, \ldots, n.$$

By Lemma A.2, the eigenvalues of the difference matrix $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{I} + (\sigma^2\rho - \sigma^{*2}\rho^*)\mathbf{T}$ are

$$\lambda_i = (\sigma^2 - \sigma^{*2}) + 2\left|\sigma^2\rho - \sigma^{*2}\rho^*\right| \cos\frac{i\pi}{n+1}, i = 1, \ldots, n.$$

Given any (σ^2, ρ) and (σ^{*2}, ρ^*), with $(\sigma^2, \rho) \neq (\sigma^{*2}, \rho^*)$, the number of zero λ_i's is at most one. Hence, the rank of the difference matrix is greater than or equal to $n-1$. Model identifiability then follows from Theorem 4.2. □

Theorem 4.3

Proof. Let $\mathbf{G_Z} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}$ and let $\mathbf{D} \equiv \mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^* = \left(\begin{smallmatrix} d_1 & d_2 \\ d_2 & d_3 \end{smallmatrix}\right)$. Direct calculation yields

$$\mathbf{G}_\mathbf{Z}'\mathbf{D}\mathbf{G}_\mathbf{Z} = \frac{1}{(z-v)^2}\begin{pmatrix} z^2d_1 - 2vzd_2 + v^2d_3 & -zd_1 + (z+v)d_2 - vd_3 \\ -zd_1 + (z+v)d_2 - vd_3 & d_1 - 2d_2 + d_3 \end{pmatrix}. \tag{12}$$

We define a VC structured $\mathbf{\Sigma}_u$ and $\mathbf{\Sigma}_\epsilon$ as $\mathrm{diag}\{\sigma_{u1}^2, \sigma_{u2}^2\}$ and $\mathrm{diag}\{\sigma_1^2, \sigma_2^2\}$ respectively. We define elements of MI structured $\mathbf{\Sigma}_u$ and $\mathbf{\Sigma}_\epsilon$ as σ_u^2 and σ^2 respectively.

The VC or MI structure of $\mathbf{\Sigma}_\epsilon$ implies $d_2 = 0$, and so (12) simplifies to

$$\frac{1}{(z-v)^2}\begin{pmatrix} z^2d_1 + v^2d_3 & -zd_1 - vd_3 \\ -zd_1 - vd_3 & d_1 + d_3 \end{pmatrix}. \tag{13}$$

To prove (a) or (b), we notice that a MI structured $\mathbf{\Sigma}_\epsilon$ has $d_1 = d_3$. If $z = -v$, (13) further simplifies to

$$\frac{1}{2z^2}\begin{pmatrix} z^2d_1 & 0 \\ 0 & d_1 \end{pmatrix}.$$

Given σ^2, we define $\sigma^{*2} = r\sigma^2$. To prove (a), we let $0 < r < 1$. It follows $d_1 > 0$ and $\mathbf{\Sigma}_u^*$ is VC structured. To prove (b), we further require $z^2 = 1$ and it follows that $\mathbf{\Sigma}_u^*$ is MI structured.

To prove (c), we first define $\sigma_1^{*2} = r\sigma_1^2$. We then define $\sigma_2^{*2} = s\sigma_2^2$, where $s = (z\sigma_1^2)(1-r)/(v\sigma_2^2) + 1$. We will show at last the existence of r such that σ_1^{*2} and σ_2^{*2} are positive. By this definition, we have $zd_1 + vd_3 = 0$ and thus the off-diagonal elements of (13) are zero. Then (13) further simplifies to

$$\frac{1}{(z-v)^2}\begin{pmatrix} z(z-v)d_1 & 0 \\ 0 & \frac{1}{v}(v-z)d_1 \end{pmatrix}.$$

In order for $\mathbf{\Sigma}_u^*$ to have a MI structure, we require $z(z-v) = (v-z)/v$ which is equivalent to $zv = -1$. As a result, the diagonal elements of the above matrix are $(z^2+1)d_1/(z-v)^2$. Clearly if $r < 1$, $\mathbf{\Sigma}_u^*$ then is MI structured. We let $\max\{0, 1 - \sigma_2^2/(z^2\sigma_1^2)\} < r < 1$ which ensures that $d_1 > 0$ and $s > 0$. □

Theorem 4.4

Proof. To prove (a), we notice if $\mathbf{\Sigma}_u$ is known, then $\mathbf{Z}\mathbf{\Sigma}_u\mathbf{Z}'$ is known, and so, by examination of (2), $\mathbf{\Sigma}_\epsilon$ is completely determined. Similarly, to prove (b), if $\mathbf{\Sigma}_\epsilon$ is known, then $\mathbf{\Sigma}_\epsilon^* = \mathbf{\Sigma}_\epsilon$ and it follows that $\mathbf{Z}\mathbf{\Sigma}_u\mathbf{Z}' = \mathbf{Z}\mathbf{\Sigma}_u^*\mathbf{Z}'$. We have $\mathbf{\Sigma}_u = \mathbf{\Sigma}_u^*$ since $\mathbf{Z}$ is of full column rank.

By assumption, we have $\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' = \mathbf{K}\boldsymbol{\Sigma}_\epsilon$ and $\mathbf{Z}\boldsymbol{\Sigma}_u^*\mathbf{Z}' = \mathbf{K}\boldsymbol{\Sigma}_\epsilon^*$. Substituting these expressions into (2) yields

$$\mathbf{K}(\boldsymbol{\Sigma}_\epsilon^* - \boldsymbol{\Sigma}_\epsilon) = \boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*,$$

that is,

$$(\mathbf{K} + \mathbf{I})(\boldsymbol{\Sigma}_\epsilon^* - \boldsymbol{\Sigma}_\epsilon) = \mathbf{0}.$$

Since $\mathbf{K} + \mathbf{I}$ is of full rank, we must have $\boldsymbol{\Sigma}_\epsilon^* = \boldsymbol{\Sigma}_\epsilon$ which is condition (a). □

Theorem 4.5

Proof. We notice each individual model (1) shares a common parameter $\boldsymbol{\Sigma}_u$. If one individual model uniquely determines $\boldsymbol{\Sigma}_u$ and its $\boldsymbol{\Sigma}_{\epsilon i}$, the identified $\boldsymbol{\Sigma}_u$ will then help to identify $\boldsymbol{\Sigma}_{\epsilon j}$ for each other j-th individual model by Theorem 4.4. Therefore, the joint model is identifiable. On the other hand, if the joint model is identifiable, then clearly each individual model is identifiable. □

References

[1] J. Cheng, L. J. Edwards, M. M. Maldonado-Molina, K. A. Komro, and K. E. Muller. Real longitudinal data analysis for real people: Building a good enough mixed model. *Statistics in Medicine*, 29:504–520, 2010. MR2751785

[2] E. Demidenko. *Mixed Models: Theory and Applications*. John Wiley and Sons, 2004. MR2077875

[3] W. A. Fuller. *Measurement Error Models*. Wiley-Interscience, 1987. MR0898653

[4] F. A. Graybill. *Matrices with Applications in Statistics*. Wadsorth, Inc, second edition edition, 1983. MR0682581

[5] R. I. Jennrich and M. D. Schluchter. Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, 42:805–820, 1986. MR0872961

[6] J. Jiang. *Linear and generalized linear mixed models and their applications*. Springer, 2010. MR2308058

[7] N. M. Laird and J. H. Ware. Random-effects models for longitudinal data. *Biometrics*, 38:963–974, 1982.

[8] M. Lindstrom and D. Bates. Newton-raphson and em algorithms for linear mixed effects models for repeated measures data. *Journal of American Statistical Association*, 83:1014–1022, 1988. MR0997577

[9] R. C. Littell, P. R. Henry, and C. B. Ammerman. Statistical analysis of repeated measures data using sas procedures. *Journal of Animal Science*, 76:1216–1231, 1998.

[10] R. C. Littell, G. A. Milliken, W. W. Stroup, R. S. Wolfinger, and O. Schabenberber. *SAS for Mixed Models*. SAS Publishing, 2nd edition edition, 2006.

[11] R. C. Littell, J. Pendergast, and R. Natarajan. Modelling covariance structure in the analysis of repeated measures data. *Statistics in Medicine*, 19:1793–1819, 2000.

[12] C. E. McCulloch, S. R. Searle, and J. M. Neuhaus. *Generalized, Linear, and Mixed Models*. Wiley, 2nd edition edition, 2008. MR2431553

[13] J. Pinheiro and D. Bates. *Mixed-Effects Models in S and S-PLUS*. Springer, 2009.

[14] B. Schaajle, J. Zhang, S. G. Pantual, and K. H. Pollock. Analysis of repeated-measurements data from randomized block experiments. *Biometrics*, 47:813–824, 1991.

[15] S. R. Searle, G. Casella, and C. E. McCulloch. *Variance Components*. Wiley-Interscience, 2006. MR2298115

[16] G. Verbeke and G. Molenberghs. *Linear Mixed Models in Practice: A SAS-Oriented Approach*. Springer, 1997.

[17] G. Verbeke and G. Molenberghs. *Linear Mixed Models for Longitudinal Data*. Springer, 2009. MR2723365

[18] B. West, K. B. Welch, and A. T. Galecki. *Linear Mixed Models: A Practical Guide Using Statistical Software*. Chapman and Hall/CRC, 2006.

[19] R. Wolfinger. Covariance structure selection in general mixed models. *Communications in Statistics: Simulation and Computation*, 22(4):1079–1106, 1993.