Identifiability in linear mixed effects models

Wei Wang

November 5th, 2007

Identifiability in linear mixed effects models

Wei Wang

November 5, 2007

Abstract

In mixed effects model, observations are a function of fixed and random effects and an error term. This structure determines a very specific structure for the variances and covariances of these observations. Unfortunately, the specific parameters of this variance/covariance structure might not be identifiable. Nonidentifiability can lead to complications in numerical estimation algorithms or worse, to incorrect or ambiguous inference. We study the identifiability of normal linear mixed effects models. We derive necessary and sufficient conditions of identifiability and we study identifiability in some commonly used variance-covariance structures. The results are particularly timely, given the recent interest in linear mixed effects models within the longitudinal and functional data analysis literature. With that in mind, we extend our discussion to identifiability in models for scalar responses depending on function-valued covariates.

1 Introduction

Mixed effects models have moved beyond the original simple models, becoming more complicated, containing more parameters. However, two different sets of parameters producing the same covariance matrix for the observations may cause problems for parameter estimation and for inference. In principle, parameter identifiability is the first thing that should be verified when building a model. As Demidenko (2004, p.118) states, "identifiability may be viewed as a necessary property for the adequacy of a statistical model".

We study identifiability in normal linear mixed effects models. In the next section, we define the classical mixed effects model and show that in an unrestricted form and a specific restricted form, the model is not identifiable. Sections 3 and 4 give sufficient conditions to identify parameters in various models. In Section 5 we discuss identifiability of extended models.

2 Motivation

In this section, we give the definition of nonidentifiability. Then we introduce the general unrestricted classical linear mixed effects model and give two models that are not nonidentifiable.

Definition 2.1 *Let $\mathbf{y}$ be the vector of observable random variables with distribution function $\mathcal{P}_{\boldsymbol{\theta}}$ where $\boldsymbol{\theta}$ is in the parameter space $\boldsymbol{\Theta}$. This probability model for $\mathbf{y}$ is not identifiable if and only if there exist $\boldsymbol{\theta}, \boldsymbol{\theta}^* \in \boldsymbol{\Theta}$, with $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$ and $\mathcal{P}_{\boldsymbol{\theta}} = \mathcal{P}_{\boldsymbol{\theta}^*}$.*

Throughout the paper, we assume all random variables follow normal distributions. As the normal distribution is uniquely characterized by its first two moments, identifiability of a normal distribution function then reduces to the identifiability of its mean vector and covariance matrix (Demidenko, 2004, *Proposition 10*, p.118).

In the standard linear mixed effects model, $\mathbf{y}$ is the observable random vector of length n and $\mathbf{X}$ and $\mathbf{Z}$ are known, non-random design matrices with dimensions $n \times p$, $n > p$ and $n \times q$, $n > q$ respectively. We assume throughout that both $\mathbf{X}$ and $\mathbf{Z}$ have full column rank. Then

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

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

The random effects vector $\mathbf{u}$ and the error vector $\boldsymbol{\epsilon}$ are unobservable. This model has been studied and applied by, for instance, McCulloch and Searle (2001).

Unknown parameters in the model are $(\boldsymbol{\beta}, \boldsymbol{\theta})$, where $\boldsymbol{\beta} \in \mathbf{B} \subseteq \Re^p$, $\boldsymbol{\theta} = (\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u) \in \tilde{\boldsymbol{\Theta}} \subseteq \boldsymbol{\Theta} = $ the set of all $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u)$ with $\boldsymbol{\Sigma}_\epsilon$, $n \times n$, and $\boldsymbol{\Sigma}_u$, $q \times q$, both symmetric and positive definite. Throughout the paper, we assume that $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ do not have common elements, i.e. we assume that $(\boldsymbol{\beta}, \boldsymbol{\theta}) \in \mathbf{B} \otimes \tilde{\boldsymbol{\Theta}}$. We also assume that $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_u$ do not have common elements. That is, we sometimes assume that $\tilde{\boldsymbol{\Theta}} = \tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}} \otimes \tilde{\boldsymbol{\Theta}}_{\mathbf{u}}$ where $\tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}} \subseteq \boldsymbol{\Theta}_{\boldsymbol{\epsilon}}$ and $\tilde{\boldsymbol{\Theta}}_{\mathbf{u}} \subseteq \boldsymbol{\Theta}_{\mathbf{u}}$, and $\boldsymbol{\Theta}_{\boldsymbol{\epsilon}}$ contains all $n \times n$ positive definite symmetric matrices and $\boldsymbol{\Theta}_{\mathbf{u}}$ contains all $q \times q$ positive definite symmetric matrices. We take as our operating definition of identifiability the ability to identify the parameters $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_u$.

The linear mixed effects model has become popular in the analysis of longitudinal and functional data. For instance, the response of individual i at time t can be modelled as $\sum \alpha_j \phi_j(t) + \sum u_{ik} \psi_k(t)$ plus error, where the α_j's are fixed population effects and the u_{ik}'s are individual-specific random effects. The inclusion of random effects allows us to realistically model covariance within an individual. Taking ϕ_j's and ψ_k's equal to B-splines allows us to flexibly model a wide variety of responses. In addition, if we allow some of the α_j's to be random but with a very specific covariance structure, then the resulting best linear unbiased predictors of the α_j's are computationally equivalent to the parameter estimates in a smoothing spline fit or a penalized B-spline fit. See Verbyla (1999) or Ruppert, Wand and Carroll (2003).

The parameter vector $\boldsymbol{\beta}$ is always identifiable since $\mathrm{E}\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$ and $\mathbf{X}$ is of full column rank. So to study identifiability in the normal model, it suffices to study the covariance matrix of $\mathbf{y}$, $\boldsymbol{\Sigma}_y = \mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon$, to see when $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_u$ are identifiable. Thus, nonidentifiability of model (1) with parameter space $\mathbf{B} \otimes \tilde{\boldsymbol{\Theta}}$ is equivalent to the existence of $\boldsymbol{\theta}, \boldsymbol{\theta}^* \in \tilde{\boldsymbol{\Theta}}$, with $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$ giving the same $\boldsymbol{\Sigma}_y$, i.e.

$$\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon = \mathbf{Z}\boldsymbol{\Sigma}_u^*\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon^*. \tag{2}$$

In Example 2.1 below, we show that the unrestricted model where $\tilde{\boldsymbol{\Theta}} = \boldsymbol{\Theta}$ is nonidentifiable. A similar argument shows that the restricted model in Example 2.2 below is not identifiable. Example 2.2 uses the covariance structure assumed for the random effects in the penalized Bspline method. Thus, when using penalized Bsplines, one cannot assume a general form for $\boldsymbol{\Sigma}_\epsilon$.

Example 2.1

Suppose we place no further restrictions on $\boldsymbol{\Theta}$. Then the model is not identifiable. To see this, let $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u) \in \boldsymbol{\Theta}$ and $0 < a < 1$. Then $\boldsymbol{\Sigma}_u^* = (1-a)\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_\epsilon{}^* = \boldsymbol{\Sigma}_\epsilon + a\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}'$. So $\boldsymbol{\Sigma}_\epsilon{}^*$ and $\boldsymbol{\Sigma}_u^*$ satisfy (2) and $(\boldsymbol{\Sigma}_\epsilon{}^*, \boldsymbol{\Sigma}_u^*) \in \boldsymbol{\Theta}$.

Example 2.2

Suppose that $\boldsymbol{\Theta} = \boldsymbol{\Theta}_{\boldsymbol{\epsilon}} \otimes \tilde{\boldsymbol{\Theta}}_{\mathbf{u}}$ where $\tilde{\boldsymbol{\Theta}}_{\mathbf{u}}$ contains all matrices of the form $\sigma^2 \mathbf{R}$ where $\mathbf{R}$ is positive definite and known, and $\sigma^2 > 0$. To see that this model is not identifiable, use the same argument as in Example 2.1 and note that the constructed $\boldsymbol{\Sigma}_u^*$ is in $\tilde{\boldsymbol{\Theta}}_{\mathbf{u}}$ if $\boldsymbol{\Sigma}_u^*$ is in $\tilde{\boldsymbol{\Theta}}_{\mathbf{u}}$.

In practice, one usually assumes a more specific structure for $\boldsymbol{\Sigma}_\epsilon$, such as $\boldsymbol{\Sigma}_\epsilon = \sigma_\epsilon^2 \mathbf{I}$. Restrictions may lead to identifiability, and such restrictions and their effects on identifiability will be discussed in the next two sections.

3 Simple sufficient conditions of identifiability

In this section, we find sufficient conditions of identifiability of model (1) assuming $\tilde{\boldsymbol{\Theta}} = \boldsymbol{\Theta}$.

A further examination of (2) gives us the following sufficient conditions. Clearly, if $\boldsymbol{\Sigma}_u$ is known, then $\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}'$ is known, and so $\boldsymbol{\Sigma}_\epsilon$ is completely determined.

If $\boldsymbol{\Sigma}_\epsilon$ is known, then the model is identifiable. To see this, consider (2) with $\boldsymbol{\Sigma}_\epsilon = \boldsymbol{\Sigma}_\epsilon{}^*$. It follows that $\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' = \mathbf{Z}\boldsymbol{\Sigma}_u^*\mathbf{Z}'$ and so $\boldsymbol{\Sigma}_u = \boldsymbol{\Sigma}_u^*$ since $\mathbf{Z}$ is of full column rank.

If $\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}'\boldsymbol{\Sigma}_\epsilon^{-1} = \mathbf{K}$, where $\mathbf{K}$ is known and $\mathbf{K} + \mathbf{I}$ is of full column rank, then the model is identifiable. Suppose by way of contradiction the model is not identifiable. Then (2) holds for $(\boldsymbol{\Sigma}_u, \boldsymbol{\Sigma}_\epsilon) \neq (\boldsymbol{\Sigma}_u^*, \boldsymbol{\Sigma}_\epsilon^*)$ both in $\tilde{\boldsymbol{\Theta}}$ with, by assumption, $\mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' = \mathbf{K}\boldsymbol{\Sigma}_\epsilon$ and $\mathbf{Z}\boldsymbol{\Sigma}_u^*\mathbf{Z}' = \mathbf{K}\boldsymbol{\Sigma}_\epsilon^*$. Substituting these expressions into (2) yields

$$\mathbf{K}\boldsymbol{\Sigma}_\epsilon + \boldsymbol{\Sigma}_\epsilon = \mathbf{K}\boldsymbol{\Sigma}_\epsilon^* + \boldsymbol{\Sigma}_\epsilon^*,$$

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 $\boldsymbol{\Sigma}_\epsilon = \boldsymbol{\Sigma}_\epsilon^*$. But, as shown in the previous paragraph, this implies that $\boldsymbol{\Sigma}_u = \boldsymbol{\Sigma}_u^*$.

The last condition is similar to a common condition for identifiabililty in simple linear regression models with measurement errors. The model assumes

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \sim (0, \sigma_\epsilon^2),$$

where x_i is observed with error having variance σ_u^2. The response y_i then has variance $\sigma_u^2 + \sigma_\epsilon^2$. One of the common conditions of model identifiability is to assume the ratio $\sigma_u^2/\sigma_\epsilon^2$ is known. The inverse $\boldsymbol{\Sigma}_\epsilon^{-1}$ appearing in our last condition could be viewed as multivariate version of "denominator".

If there are any supplementary data, we may then be able to find an estimate of $\boldsymbol{\Sigma}_u$, $\boldsymbol{\Sigma}_\epsilon$ or $\mathbf{K}$ and we can treat this estimate as the true value. The sufficient conditions for identifiability can then be satisfied.

4 Sufficient conditions of identifiability for a structured $\boldsymbol{\Sigma}_\epsilon$

As we observed from Examples 2.1 and 2.2, the model is not identifiable even if we restrict $\boldsymbol{\Sigma}_u$ to be a scalar multiple of a known matrix. In this section, we study the effect of putting restrictions on $\boldsymbol{\Sigma}_\epsilon$. In Theorem 4.1 below, we give a necessary and sufficient condition of nonidentifiability, a condition that relies mainly on the design matrix $\mathbf{Z}$ via $\mathbf{H_Z} = \mathbf{Z}(\mathbf{Z'Z})^{-1}\mathbf{Z'}$. The theorem leads to four corollaries: Corollaries 4.1 and 4.2 give necessary and sufficient conditions for identifiability when $\boldsymbol{\Sigma}_\epsilon$ arises from an exchangeable covariance structure or is diagonal. Corollary 4.3 states an easily checked condition on $\boldsymbol{\Sigma}_\epsilon$ that guarantees identifiability of the model. That corollary is then applied to two commonly used error structures. Using Corollary 4.4, we can generalize a known identifiability result, giving a shorter proof under weaker conditions.

Theorem 4.1 *Let* $\tilde{\boldsymbol{\Theta}} \subseteq \boldsymbol{\Theta}$ *and define* $\mathbf{H_Z} = \mathbf{Z}(\mathbf{Z'Z})^{-1}\mathbf{Z'}$. *Then model (1) with parameter space* $\mathbf{B} \otimes \tilde{\boldsymbol{\Theta}}$ *is nonidentifiable if and only if there exist* $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u) \in \tilde{\boldsymbol{\Theta}}$ *and* $(\boldsymbol{\Sigma}_\epsilon^*, \boldsymbol{\Sigma}_u^*) \in \tilde{\boldsymbol{\Theta}}$, *with* $\boldsymbol{\Sigma}_\epsilon^* \neq \boldsymbol{\Sigma}_\epsilon$ *such that*

$$\mathbf{H_Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right] = \boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*, \tag{3}$$

and

$$\boldsymbol{\Sigma}_u^* = \boldsymbol{\Sigma}_u + (\mathbf{Z'Z})^{-1}\mathbf{Z'}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*\right]\mathbf{Z}(\mathbf{Z'Z})^{-1}. \tag{4}$$

Proof:

Nonidentifiability of the model is equivalent to the existence of $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u)$ and

8

$(\boldsymbol{\Sigma}_\epsilon{}^*, \boldsymbol{\Sigma}_u^*)$ in $\tilde{\boldsymbol{\Theta}}$, not equal, satisfying (2). Note that this is equivalent to having $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u)$ and $(\boldsymbol{\Sigma}_\epsilon{}^*, \boldsymbol{\Sigma}_u^*)$ in $\tilde{\boldsymbol{\Theta}}$ with $\boldsymbol{\Sigma}_\epsilon{}^* \neq \boldsymbol{\Sigma}_\epsilon$ satisfying

$$\mathbf{Z}(\boldsymbol{\Sigma}_u - \boldsymbol{\Sigma}_u^*)\mathbf{Z}' = \boldsymbol{\Sigma}_\epsilon{}^* - \boldsymbol{\Sigma}_\epsilon. \tag{5}$$

Suppose the model is nonidentifiable. We premultiply (5) by $\mathbf{Z}'$, postmultiply it by $\mathbf{Z}$ and then pre- and postmultiply by $(\mathbf{Z}'\mathbf{Z})^{-1}$ to get

$$\boldsymbol{\Sigma}_u - \boldsymbol{\Sigma}_u^* = (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\left[\boldsymbol{\Sigma}_\epsilon{}^* - \boldsymbol{\Sigma}_\epsilon\right]\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}. \tag{6}$$

This gives (4). To derive (3), premultiply (6) by $\mathbf{Z}$, postmultiply (6) by $\mathbf{Z}'$ to get

$$\mathbf{Z}(\boldsymbol{\Sigma}_u - \boldsymbol{\Sigma}_u^*)\mathbf{Z}' = \mathbf{H_Z}\left[\boldsymbol{\Sigma}_\epsilon{}^* - \boldsymbol{\Sigma}_\epsilon\right]\mathbf{H_Z} \tag{7}$$

which, by (5), is the same as

$$\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^* = \mathbf{H_Z}\left[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*\right]\mathbf{H_Z}. \tag{8}$$

Premultiplying (8) by the idempotent matrix $\mathbf{H_Z}$ gives

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

Substituting (8) into the right side of the above yields (3).

To prove the converse, we want to show that (3) and (4) lead to (5). It is clear from (4) that (7) holds. If we can show that (8) holds then we are done since substituting (8) into the right side of (7) yields (5). To show (8), from (3) and the symmetry of $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*$, we see that

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

Premultiplying the above identity by the idempotent matrix $\mathbf{H_Z}$ gives $\mathbf{H_Z}[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*] = \mathbf{H_Z}[\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*]\mathbf{H_Z}$. Substituting (3) for the left side of the equation, we see that (8) holds. $\square$

The proofs of the next two corollaries are in the appendix, Section 6.

Corollary 4.1 *Let $\mathbf{1}$ be an n-vector with each element being one. Suppose that the distribution of $\epsilon_1, \ldots, \epsilon_n$ is exchangeable, that is, the covariance matrix of $\boldsymbol{\epsilon}$ is of the form $\sigma^2\left[(1-\rho)\mathbf{I} + \rho\mathbf{J}\right]$ where $\mathbf{J} = \mathbf{1}\mathbf{1}'$. Let $\tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}} = \{\boldsymbol{\Sigma}_\epsilon = \sigma^2\left[(1-\rho)\mathbf{I} + \rho\mathbf{J}\right], \ \sigma^2 > 0, \ -1/(n-1) < \rho < 1\}$. Suppose the matrix $\mathbf{Z}$ satisfies $\mathbf{1}'\mathbf{Z} \neq \mathbf{0}$ and $\mathrm{rank}(\mathbf{Z}) = q$ with $1 \le q < n-1$. Suppose the parameter space is $\mathbf{B} \otimes \tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}} \otimes \boldsymbol{\Theta}_{\mathbf{u}}$. Then model (1) is identifiable if and only if $\mathbf{H}_{\mathbf{Z}}\mathbf{J} \neq \mathbf{J}$.*

Comments. The condition $\mathbf{H}_{\mathbf{Z}}\mathbf{J} = \mathbf{J}$ means the sum of the elements of each row of $\mathbf{H}_{\mathbf{Z}}$ is equal to one, and this is an easy condition to check. For the case that $q = 1$, i.e. $\mathbf{Z}$ is a column vector $(z_1, \ldots, z_n)'$ where $\sum z_i \neq 0$,

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

The model is identifiable if and only if $\mathbf{Z}$ is not a constant vector.

When $q = 2$, suppose we have the usual simple linear regression model with centered covariates:

$$
\mathbf{Z} = \begin{pmatrix} 1 & z_1 \\ \vdots & \vdots \\ 1 & z_n \end{pmatrix}, \quad \text{with} \ \ \sum z_i = 0. \tag{9}
$$

Then

$$
\mathbf{H}_{\mathbf{Z}} = \begin{pmatrix} \frac{1}{n} + \frac{z_1^2}{s_z^2} & \frac{1}{n} + \frac{z_1 z_2}{s_z^2} & \cdots & \frac{1}{n} + \frac{z_1 z_n}{s_z^2} \\ \vdots & & & \vdots \\ \frac{1}{n} + \frac{z_n z_1}{s_z^2} & \frac{1}{n} + \frac{z_n z_2}{s_z^2} & \cdots & \frac{1}{n} + \frac{z_n^2}{s_z^2} \end{pmatrix}.
$$

10

and each row of $\mathbf{H_Z}$ sums to one. Thus unfortunately, the model is not identifiable under this $\mathbf{Z}$ combined with the exchangable covariance structure.

Corollary 4.2 *Suppose that $\tilde{\mathbf{\Theta}}_{\boldsymbol{\epsilon}}$ equals the collection of all diagonal positive definite $n \times n$ matrices. Then model (1) with parameter space $\mathbf{B} \otimes \tilde{\mathbf{\Theta}}_{\boldsymbol{\epsilon}} \otimes \mathbf{\Theta_u}$ is identifiable if and only none of the diagonal elements of $\mathbf{H_Z}$ is equal to one.*

Comments. Again, the condition on $\mathbf{H_Z}$ is easy to check. Consider the case $q = 1$. As we have seen, diagonal elements of $\mathbf{H_Z}$ equal $z_i^2 / \sum z_j^2$, $i = 1, \ldots, n$. Therefore, the model is identifiable if and only if $\mathbf{Z}$ does not have $n - 1$ zero elements. Consider $q = 2$ with $\mathbf{Z}$ as in (9). The model is identifiable provided, for all i, $(1/n) + z_i^2 / \sum z_j^2$ doesn't equal 1. So typically, the model is identifiable.

The following corollary provides a sufficient condition for identifiability, a condition that can sometimes be easily checked. Consider (3). Note that the rank of $\mathbf{H_Z}(\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*)$ is at most q, since the rank of $\mathbf{H_Z}$ is q. Thus, for (3) to hold, we must be able to find some $\mathbf{\Sigma}_\epsilon$ and $\mathbf{\Sigma}_\epsilon^*$ with the rank of $\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*$ less than or equal to q. This proves the following.

Corollary 4.3 *Suppose that $\tilde{\mathbf{\Theta}} \subseteq \mathbf{\Theta}$. Then model (1) with parameter space $\mathbf{B} \times \tilde{\mathbf{\Theta}}$ is identifiable if $\mathrm{rank}(\mathbf{\Sigma}_\epsilon - \mathbf{\Sigma}_\epsilon^*) > q$ for all $\mathbf{\Sigma}_\epsilon$, $\mathbf{\Sigma}_\epsilon^*$ in the parameter space.*

Now we apply Corollary 4.3 to show model identifiability under the "multiple of a known positive definite matrix" and the "MA(1)" covariance structures respectively in next two examples.

11

Example 4.1 *Multiple of a known positive definite matrix*

Fix $\mathbf{R}$, symmetric and positive definite, and suppose $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u) \in \tilde{\boldsymbol{\Theta}}$ implies that $\boldsymbol{\Sigma}_\epsilon = \sigma^2 \mathbf{R}$, some $\sigma^2 > 0$. Consider $\boldsymbol{\Sigma}_\epsilon = \sigma^2 \mathbf{R}$ and $\boldsymbol{\Sigma}_\epsilon^* = \sigma^{*2} \mathbf{R}$, $\sigma^{*2} \neq \sigma^2$. Clearly $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2}) \mathbf{R}$ is invertible, and so is of rank n, which we have assumed is greater than q. Thus, the model is identifiable.

To show the model in Example 4.2 below is identifiable, we need the following lemma which is a result in (Graybill, 1983, p.285)

Lemma 4.1 *Let $\mathbf{T}$ be the $n \times n$ 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*

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

Example 4.2 *MA(1)*

Suppose that $n - 1 > q$. Let the components of $\boldsymbol{\epsilon}$ have the MA(1) covariance structure, i.e. of the form $\sigma^2(\mathbf{I} + \rho\mathbf{T})$. Let $\tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}} = \{\boldsymbol{\Sigma}_\epsilon = \sigma^2(\mathbf{I} + \rho\mathbf{T}), \ \sigma^2 > 0, \ |\rho| < 1/2\}$ and suppose $(\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u) \in \tilde{\boldsymbol{\Theta}}$ implies that $\boldsymbol{\Sigma}_\epsilon \in \tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}}$.

Let $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_\epsilon^* \in \tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}}$. By Lemma 4.1, the eigenvalues of the difference matrix $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{I} + (\sigma^2\rho - \sigma^{*2}\rho^*)\mathbf{T}$ 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 $n - 1$. Therefore, model (1) is identifiable under this MA(1) covariance structure.

In longitudinal or functional data analysis, usually there are N individuals with the ith individual modelled as in (1):

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

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

$$(\boldsymbol{\Sigma}_u, \boldsymbol{\Sigma}_{\epsilon i}) \in \boldsymbol{\Theta}_\mathbf{u} \otimes \tilde{\boldsymbol{\Theta}}_{\boldsymbol{\epsilon}}^i, \quad \mathbf{u}_i \text{ and } \boldsymbol{\epsilon}_i \text{ independent.} \tag{10}$$

Statistical inference is normally based on the joint model, the model of these N individuals. The following corollary gives sufficient conditions for identifiability of the joint model. The intuition behind the result is that, if we can identify $\boldsymbol{\Sigma}_u$ from one individual, then we can identify all of the $\boldsymbol{\Sigma}_{\epsilon i}$'s.

Corollary 4.4 *If an individual model (10) is identifiable, then the joint model is identifiable.*

Proof:

We notice each individual model (10) shares a common parameter, the covariance matrix $\boldsymbol{\Sigma}_u$. If one individual model uniquely determines $\boldsymbol{\Sigma}_u$ and its $\boldsymbol{\Sigma}_{\epsilon i}$, the identified $\boldsymbol{\Sigma}_u$ will then yield identifiability of all the individual $\boldsymbol{\Sigma}_{\epsilon i}$'s since, if $\mathbf{Z}_i\boldsymbol{\Sigma}_u\mathbf{Z}_i' + \boldsymbol{\epsilon}_i = \mathbf{Z}_i\boldsymbol{\Sigma}_u\mathbf{Z}_i' + \boldsymbol{\epsilon}_i^*$, clearly $\boldsymbol{\epsilon}_i = \boldsymbol{\epsilon}_i^*$. Therefore, the joint model is identifiable. $\qquad\square$

Corollary 4.4 reduces the verification of a joint model's identifiability to the individuals'. For instance, if the i-th individual model has $\mathbf{Z}_i$ of full column rank and $\boldsymbol{\Sigma}_{\epsilon i} = \sigma_\epsilon^2 \mathbf{I}_{n_i}$, where n_i is the length of $\mathbf{y}_i$, then this individual model is identifiable by Corollary 4.1 and thus so is the joint model. Note that the other individual models can still have their $\mathbf{Z}_j$'s not of full column rank.

Demidenko (2004, Chapters 2 & 3) studies the joint model but assumes the covariance matrix of $\boldsymbol{\epsilon}_i$ is $\sigma^2 \mathbf{I}_{n_i}$. Suppose each $\mathbf{Z}_i$ is of dimension $n_i \times q$. Demidenko shows that the joint model is identifiable if at least one matrix $\mathbf{Z}_i$ is of full column rank and $\sum_{i=1}^N (n_i - q) > 0$. Using our argument in the previous paragraph, the condition $\sum_{i=1}^N (n_i - q) > 0$ can be dropped. Furthermore, our result can be applied to more general $\boldsymbol{\Sigma}_\epsilon$'s.

5 Extensions

In this section, we discuss identifiability of a model in functional regression for a functional predictor $y(\cdot)$ and a scalar response w. We derive a necessary and sufficient condition of nonidentifiability for this model.

We model $y(t)$ as $\sum_j \alpha_j \phi_j(t) + \sum_k u_k \psi_k(t)$ plus error, with the ϕ_j's and ψ_k's known, the α_j's unknown and the u_k's unknown and random. The dependence of the response w on y is modelled through an unknown functional coefficient, β: $w = \beta_0 + \int \beta(t) \left[y(t) - \mathrm{E}(y(t)) \right] dt + \eta$ where η is mean 0 normal noise. Thus, for appropriately defined $\boldsymbol{\rho}$ and with $\mathbf{u} = (u_1, \ldots, u_q)'$, we can write

$$w = \beta_0 + \boldsymbol{\rho}' \mathbf{u} + \eta,$$

$$\beta_0 \in \Re, \ \boldsymbol{\rho} \in \Re^q \text{ unknown }, \ \eta \sim (0, \sigma_\eta^2) \text{ independent of } \mathbf{u}. \qquad (11)$$

The predictor y is observed at a sequence of discretized points and the observed values are contained in the vector $\mathbf{y}$, which then follows model (1). James (2002), Müller (2005) and Heckman and Wang (2007) consider this modelling of $\mathbf{y}$ and w and propose different approaches to estimate the functional coefficient, β.

We assume model (1) and (11). For our purpose of identifiability discussion here, we consider the unknown parameters to be $(\beta_0, \boldsymbol{\beta})$ and $\boldsymbol{\theta} = (\boldsymbol{\Sigma}_\epsilon, \boldsymbol{\Sigma}_u, \sigma_\eta^2, \boldsymbol{\rho})$. We suppose that $(\beta_0, \boldsymbol{\beta}) \in \mathbf{B} \subseteq \Re \otimes \Re^p$ and that $\boldsymbol{\theta} \in \tilde{\boldsymbol{\Theta}} \subseteq \boldsymbol{\Theta} = \boldsymbol{\Theta}_{\boldsymbol{\epsilon}} \otimes \boldsymbol{\Theta}_{\mathbf{u}} \otimes \Re^+ \otimes \Re^q$ where $\boldsymbol{\Theta}_{\boldsymbol{\epsilon}}$ and $\boldsymbol{\Theta}_{\mathbf{u}}$ are as before and $\Re^+$ is the set of positive real numbers.

To study identifiability, we must study the distribution of the random vector $(\mathbf{y}', w)$. We see that $\mathrm{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $\mathrm{E}(w) = \beta_0$ and $(\mathbf{y}', w)'$ has covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{Z}\boldsymbol{\Sigma}_u\mathbf{Z}' + \boldsymbol{\Sigma}_\epsilon & \mathbf{Z}\boldsymbol{\Sigma}_u\boldsymbol{\rho} \\ \boldsymbol{\rho}'\boldsymbol{\Sigma}_u\mathbf{Z}' & \boldsymbol{\rho}'\boldsymbol{\Sigma}_u\boldsymbol{\rho} + \sigma_\eta^2 \end{pmatrix}.$$

We know the parameter $\boldsymbol{\beta}$ is identifiable if the matrix $\mathbf{X}$ is of full column rank. The identifiability of β_0 is also clear. So we focus on identifying the covariance parameters $\boldsymbol{\theta}$.

Our discussion in Section 2 suggests the unrestricted model won't be identifiable. In fact, we can construct an example following Example 2.1 to show the existence of nonidentical $\boldsymbol{\theta}$ and $\boldsymbol{\theta}^*$ both in $\boldsymbol{\Theta}$ such that (2) holds and

$$\mathbf{Z}\boldsymbol{\Sigma}_u\boldsymbol{\rho} \;=\; \mathbf{Z}\boldsymbol{\Sigma}_u^*\boldsymbol{\rho}^*, \tag{12}$$

$$\boldsymbol{\rho}'\boldsymbol{\Sigma}_u\boldsymbol{\rho} + \sigma_\epsilon^2 \;=\; \boldsymbol{\rho}^{*\prime}\boldsymbol{\Sigma}_u^*\boldsymbol{\rho}^* + \sigma_\eta^{*\,2}. \tag{13}$$

Example 5.1 *Example 2.1 continued.*

Let $0 < a < 1$, $\boldsymbol{\Sigma}_u^*$ and $\boldsymbol{\Sigma}_\epsilon^*$ be as in Example 2.1, and let $\boldsymbol{\rho}^* = \boldsymbol{\rho}/(1-a)$ and $\sigma_\eta^{*2} = \sigma_\eta^2 - a\boldsymbol{\rho}'\boldsymbol{\Sigma}_u\boldsymbol{\rho}/(1-a)$. It is not hard to see (12) and (13) are satisfied. If, in addition, we restrict $a < \sigma_\eta^2/(\sigma_\eta^2 + \boldsymbol{\rho}'\boldsymbol{\Sigma}_u\boldsymbol{\rho})$, we see that σ_η^{*2} is positive.

The following theorem gives a necessary and sufficient condition of nonidentifiability.

Theorem 5.1 *Let* $\tilde{\Theta}_{\boldsymbol{\epsilon}} \subseteq \Theta_{\boldsymbol{\epsilon}}$, *let* $\tilde{\Theta} = \tilde{\Theta}_{\boldsymbol{\epsilon}} \otimes \Theta_{\mathbf{u}} \otimes \Re^+ \otimes \Re^q$. *and define* $\mathbf{H_Z} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$. *Then model (1) and (11) with parameter space* $\mathbf{B} \otimes \tilde{\Theta}$ *is nonidentifiable if and only if there exist* $(\boldsymbol{\Sigma}_{\epsilon}, \boldsymbol{\Sigma}_u, \sigma_{\eta}^2, \boldsymbol{\rho})$ *and* $(\boldsymbol{\Sigma}_{\epsilon}{}^*, \boldsymbol{\Sigma}_u^*, \sigma_{\eta}^{*2}, \boldsymbol{\rho}^*)$ *both in* $\tilde{\Theta}$, *with* $\boldsymbol{\Sigma}_{\epsilon}{}^* \neq \boldsymbol{\Sigma}_{\epsilon}$ *such that the following hold*

(a) $\mathbf{H_Z}\left[\boldsymbol{\Sigma}_{\epsilon} - \boldsymbol{\Sigma}_{\epsilon}{}^*\right] = \boldsymbol{\Sigma}_{\epsilon} - \boldsymbol{\Sigma}_{\epsilon}{}^*,$

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

(c) $\boldsymbol{\rho}^* = \boldsymbol{\Sigma}_u^{*-1}\boldsymbol{\Sigma}_u\boldsymbol{\rho},$

(d) $\sigma_{\eta}^{*2} = \sigma_{\eta}^2 + \boldsymbol{\rho}'\boldsymbol{\Sigma}_u(\boldsymbol{\Sigma}_u^{-1} - \boldsymbol{\Sigma}_u^{*-1})\boldsymbol{\Sigma}_u\boldsymbol{\rho}.$

Proof:

Nonidentifiability of the model is equivalent to the existence of nonidentical $\boldsymbol{\theta}$ and $\boldsymbol{\theta}^*$ in $\tilde{\Theta}$, satisfying (2), (12) and (13).

First suppose that (2), (12) and (13) hold for some $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$. By the argument given in Theorem 4.1, (2) implies that *(a)* and *(b)* hold. Since $\mathbf{Z}$ is of full column rank, (12) yields *(c)*. Substituting *(c)* into (13) yields *(d)*. We easily see that, if $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$ and if *(a)* through *(d)* hold, then $\boldsymbol{\Sigma}_{\epsilon} \neq \boldsymbol{\Sigma}_{\epsilon}{}^*$.

Now suppose that conditions *(a)* through *(d)* hold for some $\boldsymbol{\Sigma}_{\epsilon} \neq \boldsymbol{\Sigma}_{\epsilon}{}^*$. Again, by the argument given in Theorem 4.1, (2) holds. We easily see that (12) and (13) also hold. $\qquad\qquad\square$

6 Appendix

6.1 Proof of Corollary 4.1

To prove it, we use the following result from Graybill, 1983, p.206.

Lemma 6.1 *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* $a + (n - 1)b$.

Proof:

To prove the corollary, we use Theorem 4.1 and a proof by contradiction.

First suppose that the model is identifiable and suppose, by way of contradiction, $\mathbf{H_Z J} = \mathbf{J}$. Fix $\boldsymbol{\Sigma}_\epsilon \in \tilde{\boldsymbol{\Theta}}_\epsilon$. Let $s > 1$, $\sigma^{*2} = s\sigma^2$ and $\rho^* = (\rho - 1)/s + 1$. It is not hard to check $-1/(n-1) < \rho^* < 1$. Define $\boldsymbol{\Sigma}_\epsilon^* = \sigma^{*2}\left[(1 - \rho^*)\mathbf{I} + \rho^*\mathbf{J}\right]$. Then $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{J}$ and, since $\mathbf{H_Z J} = \mathbf{J}$, it is clear that (3) is satisfied. We now show that, for any $\boldsymbol{\Sigma}_u \in \boldsymbol{\Theta}_\mathbf{u}$, there exists $s^* > 1$ so that $\boldsymbol{\Sigma}_u^*$ defined as in (4) is positive definite whenever $1 < s < s^*$. This will show that the model is not identifiable, which contradicts our assumption. Plugging $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^* = (\sigma^2 - \sigma^{*2})\mathbf{J}$ into (4) yields

$$\boldsymbol{\Sigma}_u^* = \boldsymbol{\Sigma}_u + \sigma^2(1 - s)(\mathbf{Z'Z})^{-1}\mathbf{Z'JZ}(\mathbf{Z'Z})^{-1}. \tag{14}$$

By assumption $\mathbf{1'Z} \neq \mathbf{0}$ and $\mathbf{Z}$ is of full column rank, the matrix $(\mathbf{Z'Z})^{-1}\mathbf{Z'JZ}(\mathbf{Z'Z})^{-1}$ is non-negative definite and of rank one since $\mathbf{J} = \mathbf{11'}$. Let λ be its non-zero and thus the largest eigenvalue of $(\mathbf{Z'Z})^{-1}\mathbf{Z'JZ}(\mathbf{Z'Z})^{-1}$. Let λ_m be the smallest eigenvalue of the matrix $\boldsymbol{\Sigma}_u$, and let $s^* = \lambda_m/(\lambda\sigma^2) + 1$. For any $\mathbf{x} \in \Re^q$, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x'}\boldsymbol{\Sigma}_u^*\mathbf{x} > 0$ by the following argument.

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

Now suppose that $\mathbf{H_Z J} \neq \mathbf{J}$ and suppose, by contradiction, that the model is not identifiable. Then, by Theorem 4.1, there exist nonidentical $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_\epsilon{}^*$ satisfying (3) and, since the rank of $\mathbf{H_Z}$ is q, the rank of $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*$ is at most q. We have

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

By Lemma 6.1, the eigenvalues of $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*$ are $(\sigma^2 - \sigma^{*2}) - (\sigma^2\rho - \sigma^{*2}\rho^*)$, which is of multiplicity $n-1$ and $(\sigma^2 - \sigma^{*2}) + (n-1)(\sigma^2\rho - \sigma^{*2}\rho^*)$, of multiplicity 1. Since $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*$ is not the 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 $\mathrm{rank}(\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^*) \leq q$, 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 $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon{}^* = (\sigma^2 - \sigma^{*2})\mathbf{J}$. But plugging this into (3) yields $\mathbf{H_Z J} = \mathbf{J}$, contradicting our assumption.

$\square$

6.2 Proof of Corollary 4.2

Proof:

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

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

Thus, if $\mathbf{H_Z}[k,k] = 1$, then $\mathbf{H_Z}[k,i] = \mathbf{H_Z}[i,k] = 0$ for all $i \neq k$.

To prove the corollary, we use Theorem 4.1 and a proof by contradiction.

18

First suppose that the model is identifiable and suppose, by way of contradiction, a diagonal element of $\mathbf{H_Z}$ is equal to 1. Without loss of generality, we assume $\mathbf{H_Z}[1,1] = 1$. Then by our observation, $\mathbf{H_Z}[1,i] = \mathbf{H_Z}[i,1] = 0$ for all $i \neq 1$. Fix $\boldsymbol{\Sigma}_\epsilon = \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_n^2\} \in \tilde{\boldsymbol{\Theta}}_\epsilon$. Let σ_1^{*2} satisfy $0 < \sigma_1^{*2} < \sigma_1^2$ and define $\boldsymbol{\Sigma}_\epsilon^* = \mathrm{diag}\{\sigma_1^{*2}, \sigma_2^2, \ldots, \sigma_n^2\}$. Then $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^* = \mathrm{diag}\{\sigma_1^2 - \sigma_1^{*2}, 0, \ldots, 0\}$. It is not hard to check that (3) is satisfied. Clearly, for any $\boldsymbol{\Sigma}_u \in \boldsymbol{\Theta}_\mathbf{u}$, $\boldsymbol{\Sigma}_u^*$ defined as in (4) is also in $\boldsymbol{\Theta}_\mathbf{u}$. Thus, the model is not identifiable, which contradicts our assumption.

Now suppose that no diagonal element of $\mathbf{H_Z}$ is equal to one and suppose, by contradiction, that the model is not identifiable. Then there exists nonidentical diagonal matrices, $\boldsymbol{\Sigma}_\epsilon$ and $\boldsymbol{\Sigma}_\epsilon^*$, satisfying (3). As $\boldsymbol{\Sigma}_\epsilon \neq \boldsymbol{\Sigma}_\epsilon^*$, at least one of the diagonal elements of $\boldsymbol{\Sigma}_\epsilon - \boldsymbol{\Sigma}_\epsilon^*$ is not zero. Suppose the k-th diagonal element is not zero. By (3), the k-th diagonal element of $\mathbf{H_Z}$ must be one, contradicting our assumption. $\qquad\square$

References

[1] Demidenko, E. (2004) *Mixed Models: Theory and Applications*, John Wiley & Sons.

[2] Graybill, F. A. (1983) *Matrices with Applications in Statistics*, Second Edition, Wadsorth, Inc.

[3] Heckman, N. and Wang, W. (2007) Linear mixed models for measurement error in functional regression. Technical Report, Department of Statistics, University of British Columbia.

[4] James, G. (2002) Generalized Linear Models with Functional Predictors. *Journal of the Royal Statistical Society B* 64, 3, 411-432.

[5] McCulloch, C. & Searle, S. R. (2001) *Generalized, Linear and Mixed Models*, John Wiley & Sons.

[6] Müller, H.G. (2005) Functional modeling and classification of longitudinal data. *Scandinavian Journal Statistics* 32, 223-240.

[7] Ruppert, D., Wand, M.P. & Carroll, R.J. (2003). *Semiparametric Regression*, Cambridge University Press.

[8] Verbyla, A. P., Cullis, B. R., Kenward, M. G., & Welham, S. J. (1999). The Analysis of Designed Experiments and Longitudinal Data by Using Smoothing Splines. *Applied Statistics* 48, 269-311.