



Bilinear regression with random effects and reduced rank restrictions

Tatjana von Rosen¹ · Dietrich von Rosen^{2,3} 

Received: 24 April 2018 / Accepted: 2 June 2019
© The Author(s) 2019

Abstract

Bilinear models with three types of effects are considered: fixed effects, random effects and latent variable effects. In the literature, bilinear models with random effects and bilinear models with latent variables have been discussed but there are no results available when combining random effects and latent variables. It is shown, via appropriate vector space decompositions, how to remove the random effects so that a well-known model comprising only fixed effects and latent variables is obtained. The spaces are chosen so that the likelihood function can be factored in a convenient and interpretable way. To obtain explicit estimators, an important standardization constraint on the random effects is assumed to hold. A theorem is presented where a complete solution to the estimation problem is given.

Keywords Fixed effects · Growth curve model · Likelihood-based estimates · Random effects · Rank restrictions

Mathematics Subject Classification 62F10 · 62F30

1 Introduction

In this article, we present a multivariate linear model which incorporates repeated measurements profiles (growth curves), covariate effects, random effects and effects due to latent variables. The model can be used to analyse quite complex

✉ Dietrich von Rosen
dietrich.von.rosen@slu.se

Tatjana von Rosen
Tatjana.vonRosen@stat.su.se

¹ Department of Statistics, Stockholm University, 106 91 Stockholm, Sweden

² Department of Energy and Technology, Swedish University of Agricultural Sciences, Box 7032, 750 07 Uppsala, Sweden

³ Department of Mathematics, Linköping University, 581 83 Linköping, Sweden

data structures. We are not aware of any mathematical treatment of such a model. Focus is on obtaining explicit estimators because it is easier to understand explicit estimators than estimators derived by some algorithm. The derivation of explicit estimators also implies that it is easier to study properties of estimators and to perform model validation which is part of the statistical paradigm.

A base model here is a bilinear regression model which is often referred to as the growth curve model and was introduced by Potthoff and Roy (1964). It will be assumed that there exists covariate information (background information) which is modeled via fixed linear effects. It is also assumed that due to, for example, the sampling procedure there will be random effects which have an impact on the data, i.e., increase the variation. Further, we exploit the idea of adding latent process information to the model. When measuring many background variables, it is often the case that fewer latent processes are governing these variables. For example, if we make field trials unobserved soil characteristics can be important and it seems reasonable to think of the soil characteristics as latent variables. Another example is when measuring EEG signals on many places on the scalp, the response can be thought to be govern by a few latent variables. The latent variables in this article are taken into account by supposing rank restrictions on parameters and in our presentation a rank restriction on the mean parameters is applied. Note that sometimes in the literature, latent variables are motivating the use of random effects which is a different implementation of the concept of latent variable.

Before defining the model, some notation are introduced. Bold upper cases denote matrices: $\mathcal{C}(\mathbf{A})$ is the column vector space generated by the columns of \mathbf{A} and $\mathcal{C}(\mathbf{A})^\perp$ denotes its orthogonal complement. The orthogonal projector on $\mathcal{C}(\mathbf{A})$ is denoted \mathbf{P}_A and equals $\mathbf{P}_A = \mathbf{A}(\mathbf{A}'\mathbf{A})^- \mathbf{A}'$, where “-” denotes an arbitrary generalized inverse (g-inverse). Note that $\mathbf{I} - \mathbf{P}_A$ is a projector on $\mathcal{C}(\mathbf{A})^\perp$. The rank of \mathbf{A} is denoted $r(\mathbf{A})$. Moreover, we will often write $(\mathbf{M})()$ instead of $(\mathbf{M})(\mathbf{M})'$, where \mathbf{M} represents any matrix expression. The matrix normal distribution with mean $\boldsymbol{\mu}: p \times n$ and dispersion $\boldsymbol{\Psi} \otimes \boldsymbol{\Sigma}$ (the symbol \otimes stands for the Kronecker product) is denoted $N_{p,n}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Psi})$ for matrices of size $p \times n$ and positive semi-definite matrices $\boldsymbol{\Sigma}: p \times p$ and $\boldsymbol{\Psi}: n \times n$ (see Ohlson et al. 2013).

Now the model which will be considered is presented in detail.

Definition 1 Let

$$\mathbf{X} = \mathbf{A}\mathbf{B}_1\mathbf{C}_1 + \mathbf{B}_2\mathbf{C}_2 + \boldsymbol{\Theta}\mathbf{F} + \mathbf{U}\mathbf{Z} + \mathbf{E},$$

where $\mathbf{A}: p \times q_1$, $\mathbf{C}_1: k_1 \times n$, $\mathbf{C}_2: k_2 \times n$, $\mathbf{F}: k_3 \times n$, are all known matrices, $\mathcal{C}(\mathbf{F}') \subseteq \mathcal{C}(\mathbf{C}'_1)$, $r(\boldsymbol{\Theta}) = q_2 < \min(p, k_3)$, $\mathbf{Z}: k_4 \times n$, $p \leq k_4$, $\mathbf{Z}\mathbf{Z}' = \mathbf{I}_{k_4}$, $\mathbf{U} \sim N_{p,k_4}(\mathbf{0}, \boldsymbol{\Sigma}_u, \mathbf{I}_{k_4})$, $\mathbf{E} \sim N_{p,n}(\mathbf{0}, \boldsymbol{\Sigma}_e, \mathbf{I}_n)$, where \mathbf{U} and \mathbf{E} are independently distributed.

Note the important standardization constraint $\mathbf{Z}\mathbf{Z}' = \mathbf{I}$ which will be utilized later. The parameters which are to be estimated are $\mathbf{B}_1, \mathbf{B}_2, \boldsymbol{\Theta}, \boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_e$. Instead of $\mathcal{C}(\mathbf{F}') \subseteq \mathcal{C}(\mathbf{C}'_1)$ given in Definition 1 some other condition can be used which

follows from the derivation of the estimates in the next section. Since in this article only estimation is considered, the random effect UZ will not be predicted.

When in Definition 1 $B_2C_2 = \mathbf{0}$, $\Theta F = \mathbf{0}$ and $UZ = \mathbf{0}$, the classical growth curve model appears (see Potthoff and Roy 1964; von Rosen 2018). When $\Theta F = \mathbf{0}$ and $UZ = \mathbf{0}$, then we have the growth curve model with background information, i.e., a mixture of GMANOVA and MANOVA models; some references to this model, as well as more general models, are Chinchilli and Elswick (1985), Verbyla and Venables (1988), von Rosen (1989) and Bai and Shi (2007). If $B_2C_2 = \mathbf{0}$ and $\Theta F = \mathbf{0}$, then we have the growth curve model with random effects (see Ip et al. 2007) whereas if only $\Theta F = \mathbf{0}$ holds we refer to Yokoyama and Fujikoshi (1992) and Yokoyama (1995) where similar models are considered and where references to earlier works can be found. In these works, one puts structures on the covariance matrix which leads to somewhat different models than in this article.

2 Two examples

This section presents two examples where the proposed model in Definition 1 can be used.

Example 1 In the first example, a field experiment had been conducted where plant growth was studied under different treatment conditions. A couple of weeks after sowing, plant heights were collected weekly and in total p weeks were studied. In the study, treatments were randomly assigned to plots and each week plants were randomly chosen from plots and plant growth was measured. The experimental units were the plots which were modeled with respect to plant growth over time.

It was assumed that plant growth is linear and the basic model equaled:

$$X = AB_1C_1 + E, \tag{1}$$

where with the help of A the linear plant growth over p weeks was modeled. For example when $p = 4$,

$$A' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ t_1 & t_2 & t_3 & t_4 \end{pmatrix}.$$

The matrix C_1 in (1) is the design matrix connected to the treatments. If we would have had a univariate response or $A = I$, when comparing treatments, C_1 is the same design matrix as in ANOVA or MANOVA, respectively.

In particular, one wanted to take into account that fields are inhomogeneous objects with respect to soil characteristics since usually plant growth depends on soil characteristics. Plots were distributed over relatively large areas and, therefore, the characteristics varied among the plots. Thus, it was important to include in (1) soil characteristics as covariable:

$$X = AB_1C_1 + B_2C_2 + E. \tag{2}$$

Moreover, when measuring plant growth, it was not possible to measure all plants within a plot. Instead each week a number of randomly selected plants were measured. Thus, it was natural to include a random effect in (2) leading to the model:

$$X = AB_1C_1 + B_2C_2 + UZ + E. \quad (3)$$

Note that the condition $ZZ' = I_{k_4}$ standardized the random effects according to the number of plots used to study a specific treatment.

Weather variables were also considered which are important variables for plant growth. However, there were too many weather-related variables which can have an influence on plant growth. For example, every hour during the day/night cycle temperature, precipitation, wind speed, etc., was recorded. All these variables are more or less connected and it was difficult to directly measure their influence on plant growth. Therefore, the concept of latent variable was of interest and latent variables were implemented in our model via rank restrictions on a parameter matrix. Hence, we ended up with the model presented in Definition 1, i.e.,

$$X = AB_1C_1 + B_2C_2 + UZ + \Theta F + E.$$

When the rank of Θ equals r the interpretation is that there exist r latent processes affecting plant growth. The condition $\mathcal{C}(F') \subseteq \mathcal{C}(C_1')$ in Definition 1 means that the weather observations were taken at the same plots where plant growth was measured.

Example 2 In the second example, it will be shown that the model in Definition 1 can be used to analyze “small area” (small domain) problems. Small area estimation is an active research area in statistics with many applications. An introduction to the subject is for example given in the book by Rao and Molina (2015). The common thread for small area estimation problems is that a survey study (finite population case) has taken place and based on the survey the goal is to extract information about small domains by adding local information which was not accounted for in the comprehensive survey.

Sometimes survey samples are investigated several times. For example, suppose that there exists a national survey which collects information about some specific production from a certain type of companies. The survey samples are followed up once per year and this is ongoing for, say, 5 years. Moreover, suppose that there are 20 regions and each region consists of 10 subregions. The survey can consist of four companies from each subregion so the whole survey comprises 800 companies.

The survey produces for each subregion and year an estimate of the production of interest. The variance of its estimator is also obtained. However, in some sense these estimates are biased since they do not take care of local information (covariables). Moreover, the sample sizes are usually too small to draw firm conclusions about subareas. Therefore, the idea is to borrow strength across subareas via statistical models and background information. Let x be a vector where all survey estimates of a specific subarea variable are gathered (a specific production in our example). Moreover, if for these variables a linear model is assumed, we can write:

$$x = \beta' C_1 + \epsilon,$$

where $\beta' C_1$ models the true value of the variable of concern and the error term ϵ is normally distributed with zero mean and known dispersion V , where the dispersion is determined solely by the survey design. Since V is known, one can equivalently study $V^{-1/2}x$ and, therefore, without loss of generality it can be assumed that $V = I$. Since the observations in the survey are repeatedly measured, and if it is supposed that there exists a second-order polynomial trend over the years, which is of the same form for all subregions, the following model can be set up:

$$X = AB_1C_1 + E, \tag{4}$$

which models what happens with the companies over time.

Under the assumption of 20 regions, 10 subregions within each region, four companies within each subregion, repeated data collection for each company for 5 years and $V = I$, then $E \sim N_{5,800}(\mathbf{0}, \Sigma, I)$. In this case, for example,

$$A' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ t_1 & t_2 & t_3 & t_4 & t_5 \\ t_1^2 & t_2^2 & t_3^2 & t_4^2 & t_5^2 \end{pmatrix}.$$

When expressing the survey estimators with the help of A and B_1 , we cannot say this is always true. Therefore, a random error $U \sim N_{p,l}(\mathbf{0}, \Sigma_u, I_l)$ is introduced in (4) and the following model emerges:

$$X = AB_1C_1 + UZ + E, \tag{5}$$

where Z is related to C_1 , i.e., $\mathcal{C}(Z') \subseteq \mathcal{C}(C_1')$ and if additionally standardizing the effects $ZZ' = I$.

There usually exist two types of covariables. One type is accounted for in the survey study and another type of covariables are variables, for example, obtained from registers about the companies in the survey or there are base-line data which were available when the survey started. The effect of the second type of covariables can be modeled by B_2C_2 which included in (5) yields the model:

$$X = AB_1C_1 + B_2C_2 + UZ + E.$$

Moreover, suppose that a large cluster of socio-econometric variables exists. It is difficult to express a functional relationship between the survey estimators and the socio-econometric variables. Instead the idea of latent variables is employed leading to rank restrictions on an unknown parameter which should model the effect of these variables. Thus, we arrive again to the model presented in Definition 1:

$$X = AB_1C_1 + B_2C_2 + \Theta F + UZ + E.$$

3 Estimation

Let Q_1 and Q_2 be matrices of basis vectors such that

$$\mathcal{C}(Q_1) = \mathcal{C}(C_1' : C_2' : Z'), \quad \mathcal{C}(Q_2) = \mathcal{C}(C_1' : C_2' : Z')^\perp$$

and assume $Q_i'Q_i = I, i = 1, 2, Q_1'Q_2 = \mathbf{0}, v = r(C_1' : C_2' : Z'), v > k_4$.

A one–one transformation of the model in Definition 1, using Q_1 and Q_2 , yields:

$$XQ_1 = AB_1C_1Q_1 + B_2C_2Q_1 + \Theta FQ_1 + UZQ_1 + EQ_1, \tag{6}$$

$$XQ_2 = EQ_2. \tag{7}$$

Lemma 1 Let $V = Q_1'Z'ZQ_1: v \times v$. The matrix V is idempotent.

Proof Since $C(Z') \subseteq C(Q_1)$ and $ZZ' = I$ the lemma is established by straight forward calculations of VV . □

From Lemma 1, it follows that:

$$V = \Gamma \begin{pmatrix} I_{k_4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \Gamma',$$

where $\Gamma: v \times v$ is an orthogonal matrix. The identity in (6) is post-multiplied by Γ , leading to the model:

$$XQ_1\Gamma = AB_1C_1Q_1\Gamma + B_2C_2Q_1\Gamma + \Theta FQ_1\Gamma + UZQ_1\Gamma + EQ_1\Gamma. \tag{8}$$

However, since the dispersion

$$D[UZQ_1\Gamma] = \begin{pmatrix} I_{k_4} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \otimes \Sigma_u,$$

the model in (8) is split into two models. Let $\Gamma = (\Gamma_1 : \Gamma_2): v \times k_4, v \times (v - k_4)$. Then, we have three models which will be used when finding estimators:

$$\begin{aligned} XQ_1\Gamma_1 &= AB_1C_1Q_1\Gamma_1 + B_2C_2Q_1\Gamma_1 + \Theta FQ_1 + UZQ_1\Gamma_1 + EQ_1\Gamma_1, \\ UZQ_1\Gamma_1 &\sim N_{p,k_4}(\mathbf{0}, \Sigma_u, I_{k_4}), \quad EQ_1\Gamma_1 \sim N_{p,k_4}(\mathbf{0}, \Sigma_e, I_{k_4}), \end{aligned} \tag{9}$$

$$\begin{aligned} XQ_1\Gamma_2 &= AB_1C_1Q_1\Gamma_2 + B_2C_2Q_1\Gamma_2 + \Theta FQ_2\Gamma_2 + EQ_1\Gamma_2, \\ EQ_1\Gamma_2 &\sim N_{p,v-k_4}(\mathbf{0}, \Sigma_e, I_{v-k_4}), \end{aligned} \tag{10}$$

$$XQ_2 = EQ_2, \quad EQ_2 \sim N_{p,n-v}(\mathbf{0}, \Sigma_e, I_{n-v}). \tag{11}$$

The idea is to utilize (10) and (11) to estimate B_1, B_2, Θ and Σ_e . Thereafter, these estimators are inserted in (9) which yields simple estimation equations for obtaining Σ_u . Suppose that estimators $\hat{B}_1, \hat{B}_2, \hat{\Theta}$ and $\hat{\Sigma}_e$ have been obtained, and let

$$\begin{aligned} Y_0 &= XQ_1\Gamma_1 - \hat{A}\hat{B}_1C_1Q_1\Gamma_1 - \hat{B}_2C_2Q_1\Gamma_1 - \hat{\Theta}FQ_1, \\ \Psi &= \Sigma_u + \Sigma_e, \end{aligned}$$

i.e., under the assumption of no randomness in \hat{B}_1, \hat{B}_2 and $\hat{\Theta}$ we have a model:

$$Y_0 = \tilde{E}, \quad \tilde{E} \sim N_{p,k_4}(\mathbf{0}, \Psi, I_{k_4}).$$

Based on this model, the maximum likelihood estimator, under the assumption that $p \leq k_4$, $\hat{\Psi} = k_4^{-1} Y_0 Y_0'$, which means that a natural estimator of Σ_u is given by

$$\hat{\Sigma}_u = \frac{1}{k_4} Y_0 Y_0' - \hat{\Sigma}_e.$$

This estimator can only be used if $\hat{\Sigma}_u$ is positive definite. However, if the estimator is not positive definite, via eigenvalues of $\hat{\Sigma}_u$ the estimator can be modified.

Now we return to (10) and (11) to estimate $\hat{B}_1, \hat{B}_2, \hat{\Theta}$ and $\hat{\Sigma}_e$, and it is convenient to merge the models, i.e.,

$$X(Q_1 \Gamma_2 : Q_2) = AB_1(C_1 Q_1 \Gamma_2 : 0) + B_2(C_2 Q_1 \Gamma_2 : 0) + \Theta(FQ_1 \Gamma_2 : 0) + \tilde{E}, \quad \tilde{E} \sim N_{p, n-k_4}(0, \Sigma_e, I_{n-k_4}). \tag{12}$$

Let

$$Y = X(Q_1 \Gamma_2 : Q_2) : p \times (n - k_4), D_1 = (C_1 Q_1 \Gamma_2 : 0) : k_1 \times (n - k_4), \tag{13}$$

$$D_2 = (C_2 Q_1 \Gamma_2 : 0) : k_2 \times (n - k_4), D_3 = (FQ_1 \Gamma_2 : 0) : k_3 \times (n - k_4) \tag{14}$$

and then (12) is identical to

$$Y = AB_1 D_1 + B_2 D_2 + \Theta D_3 + \tilde{E}. \tag{15}$$

Furthermore, the likelihood function which corresponds to the model in (15) equals

$$\begin{aligned} L(B_1, B_2, \Theta, \Sigma_e) &= (2\pi)^{-\frac{1}{2}p(n-k_4)} |\Sigma_e|^{-\frac{1}{2}(n-k_4)} \\ &\quad \times e^{-\frac{1}{2}\text{tr}\{\Sigma_e^{-1}(Y-AB_1D_1-B_2D_2-\Theta D_3)0'\}}_V \\ &\leq (2\pi)^{-\frac{1}{2}p(n-k_4)} |\Sigma_e|^{-\frac{1}{2}(n-k_4)} \\ &\quad \times e^{-\frac{1}{2}\text{tr}\{\Sigma_e^{-1}(Y(I-P_{D'_2})-AB_1D_1(I-P_{D'_2})-\Theta D_3(I-P_{D'_2}))0'\}} \end{aligned} \tag{16}$$

with equality if and only if

$$B_2 D_2 = Y P_{D'_2} - AB_1 D_1 P_{D'_2} - \Theta D_3 P_{D'_2},$$

which, under some full rank conditions on D_2 , determines B_2 as a function of B_1 and Θ . To show the inequality, we have used

$$\begin{aligned} &(Y - AB_1 D_1 - B_2 D_2 - \Theta D_3)0' \\ &= (Y - AB_1 D_1 - B_2 D_2 - \Theta D_3) P_{D'_2} (Y - AB_1 D_1 - B_2 D_2 - \Theta D_3)' \\ &\quad + (Y - AB_1 D_1 - B_2 D_2 - \Theta D_3) (I - P_{D'_2}) (Y - AB_1 D_1 - B_2 D_2 - \Theta D_3)' \\ &= (Y - AB_1 D_1 - B_2 D_2 - \Theta D_3) P_{D'_2} (Y - AB_1 D_1 - B_2 D_2 - \Theta D_3)' \\ &\quad + (Y - AB_1 D_1 - \Theta D_3) (I - P_{D'_2}) (Y - AB_1 D_1 - \Theta D_3)', \end{aligned}$$

where both terms are positive semi-definite. The density in (16) corresponds to the model:

$$Y(I - P_{D'_2}) = AB_1D_1(I - P_{D'_2}) + \Theta D_3(I - P_{D'_2}) + E, E \sim N_{p,n-k_4}(\mathbf{0}, \Sigma_e, I_{n-k_4}). \tag{17}$$

Thus, we have a model which was treated by von Rosen and von Rosen (2017) and any further calculations are not necessary. For notational conveniences, we write the model in (17):

$$Y_1 = AB_1\tilde{D}_1 + \Theta\tilde{D}_3 + E, \quad E \sim N_{p,n-k_4}(\mathbf{0}, \Sigma_e, I_{n-k_4}), \tag{18}$$

$\tilde{D}_1 = D_1(I - P_{D'_2})$ and $\tilde{D}_3 = D_3(I - P_{D'_2})$. Note that $\mathcal{C}(\tilde{D}_3) \subseteq \mathcal{C}(\tilde{D}_1)$ because $\mathcal{C}(F') \subseteq \mathcal{C}(C'_1)$ which is essential for being able to obtain explicit estimators (see von Rosen 1989).

The following theorem presents the estimators of the parameters in the model given by Definition 1.

Theorem 1 *Let the model be as in Definition 1 and put $v = \dim \mathcal{C}(C'_1 : C'_2 : Z')$. Define Θ_1 and Θ_2 via $\Theta = \Theta_1\Theta_2$, where $\Theta_1: p \times r(\Theta)$ and $\Theta_2: r(\Theta) \times k_3$. Moreover, the matrices D_1, D_2 , and D_3 are given in (13) and (14), and Y_1, \tilde{D}_1 and \tilde{D}_3 are identified by comparing (17) and (18). Then,*

- (i) $\hat{\Theta}_1 = H(H'H)^{-1}\tilde{F}_1$, where \tilde{F} consists of the eigenvectors corresponding to the eigenvalues of the positive definite matrix

$$I_{p-r(A)} - H'Y_1P_{\tilde{D}'_3}R^{-1}P_{\tilde{D}'_3}Y_1'H,$$

where

$$R = I_v + P_{\tilde{D}'_3}Y_1'HH'Y_1P_{\tilde{D}'_3}$$

and H is defined by

$$T_1'S_2^{-1}T_1 = HH',$$

where

$$\begin{aligned} S_1 &= Y_1(I - P_{\tilde{D}'_1})Y_1', \\ T_1 &= I - A(A'S_1^{-1}A)^{-1}A'S_1^{-1}, \\ S_2 &= S_1 + T_1Y_1(P_{\tilde{D}'_1} - P_{\tilde{D}'_3})Y_1'T_1'; \end{aligned}$$

- (ii) If $\mathcal{C}(A) \cap \mathcal{C}(\Theta) = \{\mathbf{0}\}$,

$$\hat{\Theta}\tilde{D}_3 = \hat{\Theta}_1\hat{\Theta}'_1T_1'S_2^{-1}Y_1P_{\tilde{D}'_3};$$

If additionally $r(\tilde{\mathbf{D}}_3) = k_3$,

$$\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_1 \hat{\boldsymbol{\theta}}_1' \mathbf{T}'_1 \mathbf{S}_2^{-1} \mathbf{Y}_1 \tilde{\mathbf{D}}_3' (\tilde{\mathbf{D}}_3 \tilde{\mathbf{D}}_3')^{-1};$$

(iii) If $\mathcal{C}(\mathbf{A}) \cap \mathcal{C}(\boldsymbol{\theta}) = \{\mathbf{0}\}$,

$$\mathbf{A} \hat{\mathbf{B}}_1 \tilde{\mathbf{D}}_1 = \mathbf{A} (\mathbf{A}' \mathbf{S}_1^{-1} \mathbf{A})^{-1} \mathbf{A}' \mathbf{S}_1^{-1} (\mathbf{Y}_1 \mathbf{P}_{\tilde{\mathbf{D}}_1}' - \hat{\boldsymbol{\theta}} \tilde{\mathbf{D}}_3);$$

If additionally $r(\tilde{\mathbf{D}}_1) = k_1$ and $r(\mathbf{A}) = q_1$,

$$\hat{\mathbf{B}}_1 = (\mathbf{A}' \mathbf{S}_1^{-1} \mathbf{A})^{-1} \mathbf{A}' \mathbf{S}_1^{-1} (\mathbf{Y}_1 \tilde{\mathbf{D}}_1' (\tilde{\mathbf{D}}_1 \tilde{\mathbf{D}}_1')^{-1} - \hat{\boldsymbol{\theta}} \tilde{\mathbf{D}}_3' \tilde{\mathbf{D}}_1' (\tilde{\mathbf{D}}_1 \tilde{\mathbf{D}}_1')^{-1});$$

(iv)

$$\hat{\mathbf{B}}_2 \mathbf{D}_2 = \mathbf{Y} \mathbf{P}_{\mathbf{D}_2}' - \hat{\mathbf{A}} \hat{\mathbf{B}}_1 \mathbf{D}_1 \mathbf{P}_{\mathbf{D}_2}' - \hat{\boldsymbol{\theta}} \mathbf{D}_3 \mathbf{P}_{\mathbf{D}_2}';$$

If additionally $r(\mathbf{D}_2) = k_2$,

$$\hat{\mathbf{B}}_2 = \mathbf{Y} \mathbf{D}_2' (\mathbf{D}_2 \mathbf{D}_2')^{-1} - \hat{\mathbf{A}} \hat{\mathbf{B}}_1 \mathbf{D}_1 \mathbf{D}_2' (\mathbf{D}_2 \mathbf{D}_2')^{-1} - \hat{\boldsymbol{\theta}} \mathbf{D}_3 \mathbf{D}_2' (\mathbf{D}_2 \mathbf{D}_2')^{-1};$$

(v)

$$(n - k_4) \hat{\boldsymbol{\Sigma}}_e = \mathbf{S}_2 + \hat{\mathbf{T}}_2 \mathbf{T}_1 \mathbf{Y}_1 \mathbf{P}_{\tilde{\mathbf{D}}_3}' \mathbf{Y}'_1 \mathbf{T}'_1 \hat{\mathbf{T}}_2'$$

where

$$\hat{\mathbf{T}}_2 = \mathbf{I}_p - \mathbf{T}_1 \hat{\boldsymbol{\theta}}_1 (\hat{\boldsymbol{\theta}}_1' \mathbf{T}'_1 \mathbf{S}_2^{-1} \mathbf{T}_1 \hat{\boldsymbol{\theta}}_1)^{-1} \hat{\boldsymbol{\theta}}_1' \mathbf{T}'_1 \mathbf{S}_2^{-1};$$

(vi)

$$\hat{\boldsymbol{\Sigma}}_u = \frac{1}{k_4} \mathbf{Y}_0 \mathbf{Y}'_0 - \hat{\boldsymbol{\Sigma}}_e$$

Proof The proof follows from von Rosen and von Rosen (2017, Theorem 2.1) and the calculations presented in this section, in particular the model in (18) is utilized. □

A few remarks can be made: In (i), $\mathbf{P}_{\tilde{\mathbf{D}}_1}' - \mathbf{P}_{\tilde{\mathbf{D}}_3}'$ is always positive semi-definite because $\mathcal{C}(\tilde{\mathbf{D}}_3') \subseteq \mathcal{C}(\tilde{\mathbf{D}}_1')$; the condition $\mathcal{C}(\mathbf{A}) \cap \mathcal{C}(\boldsymbol{\theta}) = \{\mathbf{0}\}$ in (ii) and (iii) can always be assumed to hold because $\mathcal{C}(\boldsymbol{\theta}) = \mathcal{C}(\boldsymbol{\theta}_1)$ and any estimator $\hat{\boldsymbol{\theta}}_1$ will not be related to $\mathcal{C}(\mathbf{A})$; the estimator in (vi) is only appropriate if $\hat{\boldsymbol{\Sigma}}_u$ is positive definite; the condition $r(\tilde{\mathbf{D}}_1) = k_1$ in (iii) is fulfilled if $\mathcal{C}(\mathbf{D}'_1) \cap \mathcal{C}(\mathbf{D}'_2) = \{\mathbf{0}\}$.

Funding This research has been supported by The Swedish Foundation for Humanities and Social-Sciences (P14-0641:1), and The Swedish Natural Research Council (2017-03003).

Compliance with ethical standards

Conflict of interest On behalf of all the authors, the corresponding author states that there is no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bai, P., & Shi, L. (2007). Exact distributions of MLEs of regression coefficients in GMANOVA–MANOVA model. *Journal of Multivariate Analysis*, *98*, 1840–1852.
- Chinchilli, V., & Elswick, R. (1985). A mixture of the manova and gmanova models. *Communications in Statistics-Theory and Methods*, *14*, 3075–3089.
- Ip, W.-C., Wu, M.-X., Wang, S.-G., & Wong, H. (2007). Estimation for parameters of interest in random effects growth curve models. *Journal of Multivariate Analysis*, *98*, 317–327.
- Ohlson, M., Ahmad, R., & von Rosen, D. (2013). The multilinear normal distribution: Introduction and some basic properties. *Journal of Multivariate Analysis*, *113*, 37–47.
- Potthoff, R. F., & Roy, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, *51*, 313–326.
- Rao, J. N. K., & Molina, I. (2015). *Small Area Estimation* (2nd ed.). Hoboken: Wiley Series in Survey Methodology.
- Verbyla, A. P., & Venables, W. N. (1988). An extension of the growth curve model. *Biometrika*, *75*, 129–138.
- von Rosen, D. (1989). Maximum likelihood estimators in multivariate linear normal models. *Journal of Multivariate Analysis*, *31*, 187–200.
- von Rosen, D. (2018). *Bilinear Regression Analysis: An Introduction. Lecture Notes Statistics*, *210*, New York: Springer.
- von Rosen, T., & von Rosen, D. (2017). On estimation in some reduced rank extended growth curve models. *Mathematical Methods of Statistics*, *26*, 299–310.
- Yokoyama, T. (1995). Statistical inference on some mixed MANOVA–GMANOVA models with random effects. *Hiroshima Mathematical Journal*, *25*, 441–474.
- Yokoyama, T., & Fujikoshi, Y. (1992). Tests for random-effects covariance structures in the growth curve model with covariates. *Hiroshima Mathematical Journal*, *22*, 195–202.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.