Tatra Mt. Math. Publ. **39** (2008), 45–52

Tatra Mountains
Mathematical Publications

REMARKS ON UNBIASED ESTIMATION OF THE SUM-OF-PROFILES MODEL PARAMETERS

Ivan Žežula

ABSTRACT. The extended growth curve model (sum-of-profiles model) with two fixed effects is considered. Some results on the estimability of its parameters are presented.

1. Introduction

Let us consider a special form of the general extended growth curve model (ECGM), so called sum-of-profiles model with two components:

$$Y = X_1 B_1 Z_1 + X_2 B_2 Z_2 + e\,,$$
$$\mathrm{E}\, e = 0\,, \quad \mathrm{var}\,(\mathrm{vec}\, e) = \Sigma \otimes I$$

(dimensions of Y, e, X_i, B_i and Z_i are $n \times p, n \times p, n \times m_i, m_i \times r_i$, and $r_i \times p$, respectively).

This model was introduced by Von Rosen [6] with arbitrary number of components k. Usually X_i are ANOVA design matrices and Z_i are regression constants matrices. There are two main goals in this model: estimation of the unknown parameters and their testing. In the case of normality, Von Rosen [7] derived maximum likelihood estimators of the parameters B_i under the condition that ranges of matrices X_i are ordered:

$$\mathcal{R}\,(X_k) \subseteq \cdots \subseteq \mathcal{R}\,(X_1)\;.$$

However, these estimators are only half-explicit, computed by a complicated recurrence formula. Also, as usually, MLE of the variance matrix is not unbiased, and the bias is only to be estimated by approximate methods.

2000 Mathematics Subject Classification: Primary 62J10; Secondary 62H15.
Keywords: extended growth curve model, sum-of-profiles model, multivariate regression, unbiased estimator.
Supported by grants VEGA MŠ SR No. 1/7471/20, 1/0385/03, 1/3129/06, and Oxford University.

Therefore, we would like to derive explicit formula for unbiased estimators of B_i. Their form can reveal their geometric properties, and help us to generalize them to more complicated models. Also, they can help us to assess small sample properties of MLE.

We suppose $\mathcal{R}(X_2) \subseteq \mathcal{R}(X_1)$ throughout the paper. This assumption is inevitable for the method we use and it seems to be also a necessary condition for the existence of such estimators.

We shall denote by P_G orthogonal projector on column space $\mathcal{R}(G)$ of a matrix G and by $M_G = I - P_G$ orthogonal projector on its orthogonal complement. If the corresponding (semi)metrics is given by a p.s.d. matrix A, these projectors will be denoted by P_G^A and M_G^A.

2. Structure of unbiased estimators

When we vectorize the model, we get

$$\begin{aligned}\operatorname{vec} Y &= (Z_1' \otimes X_1) \operatorname{vec} B_1 + (Z_2' \otimes X_2) \operatorname{vec} B_2 + \operatorname{vec} e \\ &= (Z_1' \otimes X_1, Z_2' \otimes X_2) \begin{pmatrix} \operatorname{vec} B_1 \\ \operatorname{vec} B_2 \end{pmatrix} + \operatorname{vec} e \\ &\stackrel{\text{df}}{=} X\beta + e\,.\end{aligned}$$

This is a standard univariate linear model. We know that least squares estimators are unbiased and under normality have optimal properties. In order to get LSE, we need to have explicit formula for $\left(X'V^{-1}X\right)^-$, where $V = \operatorname{var} e = \Sigma \otimes I$, possibly preserving the Kronecker structure. According to Marsaglia and Styan [3], we have

$$\begin{aligned}\left(X'V^{-1}X\right)^- &= \left[\begin{pmatrix} Z_1 \otimes X_1' \\ Z_2 \otimes X_2' \end{pmatrix} \left(\Sigma^{-1} \otimes I\right) (Z_1' \otimes X_1, Z_2' \otimes X_2)\right]^- \\ &= \begin{bmatrix} Z_1\Sigma^{-1}Z_1' \otimes X_1'X_1 & Z_1\Sigma^{-1}Z_2' \otimes X_1'X_2 \\ Z_2\Sigma^{-1}Z_1' \otimes X_2'X_1 & Z_2\Sigma^{-1}Z_2' \otimes X_2'X_2 \end{bmatrix}^- \\ &= \begin{bmatrix} a + bdc & -bd \\ -dc & d \end{bmatrix},\end{aligned}$$

where

$$\begin{aligned}
a &= \left(Z_1\Sigma^{-1}Z_1'\right)^- \otimes \left(X_1'X_1\right)^-,\\
b &= \left(Z_1\Sigma^{-1}Z_1'\right)^- Z_1\Sigma^{-1}Z_2' \otimes \left(X_1'X_1\right)^- X_1'X_2\,,\\
c &= Z_2\Sigma^{-1}Z_1'\left(Z_1\Sigma^{-1}Z_1'\right)^- \otimes X_2'X_1\left(X_1'X_1\right)^-,\\
d &= \left(Z_2\Sigma^{-1}Z_2' \otimes X_2'X_2 - Z_2\Sigma^{-1}P_{Z_1'}^{\Sigma^{-1}}Z_2' \otimes X_2'P_{X_1}X_2\right)^-.
\end{aligned}$$

Main problem here is the g-inversion of the difference in d. But the assumption about ordering of $\mathcal{R}\left(X_i\right)$ implies that $P_{X_1}X_2 = X_2$ and therefore

$$d = \left(Z_2\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}Z_2' \otimes X_2'X_2\right)^- = \left(Z_2\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}Z_2'\right)^- \otimes \left(X_2'X_2\right)^- \,.$$

Using this, we easily get

$$\hat\beta = \begin{pmatrix}\operatorname{vec}\hat B_1\\ \operatorname{vec}\hat B_2\end{pmatrix} = \begin{pmatrix}K_1\\ K_2\end{pmatrix}\operatorname{vec}Y,$$

where

$$\begin{aligned}
K_1 &= \left(Z_1\Sigma^{-1}Z_1'\right)^- Z_1\Sigma^{-1} \otimes \left(X_1'X_1\right)^- X_1'\\
&\qquad - \left(Z_1\Sigma^{-1}Z_1'\right)^- Z_1\Sigma^{-1}P_{Z_2'}^{\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}} \otimes \left(X_1'X_1\right)^- X_1'P_{X_2}\,,\\
K_2 &= \left(Z_2\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}Z_2'\right)^- Z_2\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}} \otimes \left(X_2'X_2\right)^- X_2'\,.
\end{aligned}$$

De-vectorization of this expression is easy. Thus, we have unbiased LS-estimators of B_i in a closed form:

$$\begin{aligned}
\hat B_1 &= \left(X_1'X_1\right)^- X_1'Y\Sigma^{-1}Z_1'\left(Z_1\Sigma^{-1}Z_1'\right)^-\\
&\qquad - \left(X_1'X_1\right)^- X_1'P_{X_2}Y\left(P_{Z_2'}^{\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}}\right)'\Sigma^{-1}Z_1'\left(Z_1\Sigma^{-1}Z_1'\right)^-,\\
\hat B_2 &= \left(X_2'X_2\right)^- X_2'Y\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}Z_2'\left(Z_2\Sigma^{-1}M_{Z_1'}^{\Sigma^{-1}}Z_2'\right)^-.
\end{aligned}$$

In fact, speaking about unbiasedness is correct only in the case when all matrices X_1, X_2, Z_1, Z_2 have full rank, and all g-inverses turn into regular inverses. In most cases, Z_is are of full rank. However, ANOVA matrices X_is need not have full rank. If B_is are estimable then these expressions are invariant (and we could use MP-inversion). If not, these formulas produce unbiased estimators for all estimable functions of B_is.

ML estimators are much more complicated and lead only to asymptotic unbiasedness of estimable functions (see Von Rosen [8]).

It is worth noting that under normality, if Z_i are full-rank matrices, the statistic $(X_1'Y, X_2'Y, Y'Y)$ is complete sufficient statistic for all unknown parameters. Thus, in this case the estimators proposed are functions of complete sufficient statistic.

Naturally, we want to estimate Σ as well, since the estimators depend on it. Using the same method as before, we can derive that the projection matrix in the vectorized model can be expressed as

$$P^{\Sigma^{-1}\otimes I}_{(Z_1'\otimes X_1, Z_2'\otimes X_2)} = P^{\Sigma^{-1}}_{Z_1'} \otimes P_{X_1} + \left(P^{\Sigma^{-1}}_{(Z_1',Z_2')} - P^{\Sigma^{-1}}_{Z_1'}\right) \otimes P_{X_2}$$

(we made use of the formula $M^{\Sigma^{-1}}_{Z_1'} M^{\Sigma^{-1} M^{\Sigma^{-1}}_{Z_1'}}_{Z_2'} = M^{\Sigma^{-1}}_{(Z_1',Z_2')}$). This matrix is also useful when working with estimable functions, see Ž e ž u l a [10]. It follows that

$$\hat{Y} = X_1\hat{B}_1 Z_1 + X_2\hat{B}_2 Z_2 = P_{X_1} Y \left(P^{\Sigma^{-1}}_{Z_1'}\right)' + P_{X_2} Y \left(P^{\Sigma^{-1}}_{(Z_1',Z_2')} - P^{\Sigma^{-1}}_{Z_1'}\right)'.$$

It is natural to base an estimator of Σ on the expression $(Y-\hat{Y})'(Y-\hat{Y})$. In fact, the last expression—but with more complicated $\hat{Y}$—is equal to $n\hat{\Sigma}_{ML}$. Using the expression for $\hat{Y}$, after some computation we get

$$\begin{aligned}(Y-\hat{Y})'(Y-\hat{Y}) = {} & Y'M_{X_1}Y \\ & + M^{\Sigma^{-1}}_{Z_1'} Y' \left(M_{X_2} - M_{X_1}\right) Y \left(M^{\Sigma^{-1}}_{Z_1'}\right)' \\ & + M^{\Sigma^{-1}}_{(Z_1',Z_2')} Y' P_{X_2} Y \left(M^{\Sigma^{-1}}_{(Z_1',Z_2')}\right)'.\end{aligned}$$

According to G h a z a l and N e u d e c k e r [2], we obtain

$$\begin{aligned}&\mathrm{E}(Y-\hat{Y})'(Y-\hat{Y}) = \\ &\qquad = \big(n - r(X_1)\big)\Sigma + \big(r(X_1) - r(X_2)\big) M^{\Sigma^{-1}}_{Z_1'}\Sigma + r\left(X_2\right) M^{\Sigma^{-1}}_{(Z_1',Z_2')}\Sigma\,.\end{aligned}$$

We see that an unbiased estimator of Σ is

$$\hat{\Sigma} = \frac{1}{n - r\left(X_1\right)}\, Y'M_{X_1}Y,$$

which is also the REML estimator in the normal case.

It is logical that it makes use only of X_1, since X_2 contains less information, and $M_{X_1} = M_{(X_1,X_2)}$. In a less general situation, but without range ordering, V o n R o s e n [9] used matrix $M_{(X_1,X_2)}$ to get an unbiased estimator. Nevertheless, there is a situation when we can use the whole "residual squares matrix". It is the situation when the correlation structure is known, i.e., $\Sigma = \sigma^2 R$ with R

known. In such a case, the above expression for mean value immediately implies that

$$\hat{\sigma}^2 = \frac{1}{m} \operatorname{Tr}\Big[\big(Y - \hat{Y}\big)'\big(Y - \hat{Y}\big)\Big] ,$$

where

$$\begin{aligned} m = {} & \big(n - r\,(X_1)\big) \operatorname{Tr}(R) \\ & + \big(r\,(X_1) - r\,(X_2)\big) \operatorname{Tr}\left(M_{Z_1'}^{R^{-1}} R\right) \\ & + r\,(X_2) \operatorname{Tr}\left(M_{(Z_1',Z_2')}^{R^{-1}} R\right) , \end{aligned}$$

is an unbiased estimator of residual variance σ^2.

3. Generalization

Let us consider the sum-of-profiles model with k components:

$$Y = \sum_{i=1}^{k} X_i B_i Z_i + e \,,$$

with other assumptions as in the previous sections. We stress that we also suppose

$$\mathcal{R}\,(X_k) \subseteq \cdots \subseteq \mathcal{R}\,(X_1) \;.$$

Let us define

$$\begin{aligned} \hat{Y} = P_{X_1} Y \left(P_{Z_1'}^{\Sigma^{-1}}\right)' + P_{X_2} Y \left(P_{(Z_1',Z_2')}^{\Sigma^{-1}} - P_{Z_1'}^{\Sigma^{-1}}\right)' + \ldots \\ \ldots + P_{X_k} Y \left(P_{(Z_1',\ldots,Z_k')}^{\Sigma^{-1}} - P_{(Z_1',\ldots,Z_{k-1}')}^{\Sigma^{-1}}\right)' . \end{aligned}$$

Because

$$\mathcal{R}\,(Z_1') \subseteq \mathcal{R}\,(Z_1', Z_2') \subseteq \cdots \subseteq \mathcal{R}\,(Z_1', \ldots, Z_k')$$

and

$$P_{(Z_1',\ldots,Z_i')}^{\Sigma^{-1}} - P_{(Z_1',\ldots,Z_{i-1}')}^{\Sigma^{-1}} = P_{(Z_1',\ldots,Z_i')}^{\Sigma^{-1}} M_{(Z_1',\ldots,Z_{i-1}')}^{\Sigma^{-1}}$$

is the projector on $\mathcal{R}\,(Z_1', \ldots, Z_i') \cap \mathcal{R}\left(Z_1', \ldots, Z_{i-1}'\right)^{\perp}$ $\forall\, i$, we can see the geometric structure of the estimator: it adds successive projections of Y on ordered subspaces $\mathcal{R}\,(X_i)$ and difference subspaces $\mathcal{R}\,(Z_1', \ldots, Z_i') \cap \mathcal{R}\left(Z_1', \ldots, Z_{i-1}'\right)^{\perp}$. Again, if all Z_is are full-rank matrices, under normality this is a function of complete sufficient statistic.

Since

$$\begin{aligned}
&\mathrm{E}\, P_{X_i} Y \left(P^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')} - P^{\Sigma^{-1}}_{(Z_1',\dots,Z_{i-1}')} \right)' \\
&= P_{X_i} \left(\sum_{j=1}^{k} X_j B_j Z_j \right) \times \left(P^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')} M^{\Sigma^{-1}}_{(Z_1',\dots,Z_{i-1}')} \right)' \\
&= \left(\sum_{j=i}^{k} X_j B_j Z_j \right) \left(P^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')} - P^{\Sigma^{-1}}_{(Z_1',\dots,Z_{i-1}')} \right)'
\end{aligned}$$

(notice that P_{X_i} disappeared due to ordering of $\mathcal{R}\,(X_i)$), we get

$$\begin{aligned}
\mathrm{E}\, \hat{Y} = {} & X_1 B_1 Z_1 \left(P^{\Sigma^{-1}}_{Z_1'} \right)' \\
& + X_2 B_2 Z_2 \left(P^{\Sigma^{-1}}_{Z_1'} + P^{\Sigma^{-1}}_{(Z_1',Z_2')} - P^{\Sigma^{-1}}_{Z_1'} \right)' + \dots \\
& + X_k B_k Z_k \Big(P^{\Sigma^{-1}}_{Z_1'} + P^{\Sigma^{-1}}_{(Z_1',Z_2')} - P^{\Sigma^{-1}}_{Z_1'} + \dots \\
& + P^{\Sigma^{-1}}_{(Z_1',\dots,Z_k')} - P^{\Sigma^{-1}}_{(Z_1',\dots,Z_{k-1}')} \Big)' \\
= {} & \sum_{i=1}^{k} X_i B_i Z_i \Big(P^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')} \Big)' \\
= {} & \sum_{i=1}^{k} X_i B_i Z_i \,.
\end{aligned}$$

Thus, $\hat{Y}$ is an unbiased estimator of $\mathrm{E}\, Y$ in general sum-of-profiles model with Von Rosen's condition. Unfortunately, this cannot be used for the estimation of individual components. Straightforward vectorization of the general model is also not usable, because g-inverse (or inverse) of block-wise Kronecker structured matrix is not necessarily Kronecker structured.

However, we can (and should) use this estimator for the estimation of the variance matrix. Using the same way as before, we get

$$\begin{aligned}
(Y - \hat{Y})'(Y - \hat{Y}) = {} & Y' M_{X_1} Y \\
& + \sum_{i=1}^{k-1} M^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')} Y' \left(M_{X_{i+1}} - M_{X_i} \right) Y \left(M^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')} \right)' \\
& + M^{\Sigma^{-1}}_{(Z_1',\dots,Z_k')} Y' P_{X_k} Y \left(M^{\Sigma^{-1}}_{(Z_1',\dots,Z_k')} \right)'
\end{aligned}$$

and

$$\begin{aligned}\mathrm{E}(Y-\hat{Y})'(Y-\hat{Y}) =& \big(n-r(X_1)\big)\Sigma \\ &+ \sum_{i=1}^{k-1}\big(r(X_i)-r(X_{i+1})\big) M^{\Sigma^{-1}}_{(Z_1',\dots,Z_i')}\Sigma \\ &+ r\left(X_k\right) M^{\Sigma^{-1}}_{(Z_1',\dots,Z_k')}\Sigma\,.\end{aligned}$$

This implies the expected result that we should not change anything on our estimator $\hat{\Sigma}$ from the previous section. Changes required in the estimator of $\hat{\sigma}^2$ in the case $\Sigma = \sigma^2 R$ are obvious.

These results, even if they do not cover the whole scope we would like to, are substantially simpler them the ML-estimators or estimators using restricted parameter space (see, e.g., Fujikoshi [1]). The method of forming projectors on sequentially joined matrices seems to be effective in other cases, too. Whole area is open to further investigation.

Acknowledgement. The author would like to thank to Prof. D. Cox, Oxford, for his help and useful discussions while working on this paper, and for his critical remarks on the early drafts which helped to improve the quality of the paper.

REFERENCES

[1] FUJIKOSHI, Y.: *Simultaneous confidence intervals in an extended growth curve model*, Comm. Statist. Theory Methods **28** (1999), 671–682.

[2] GHAZAL, G. A.—NEUDECKER, H.: *On second-order and fourth-order moments of jointly distributed random matrices: a survey*, Linear Algebra Appl. **321** (2000), 61–93.

[3] MARSAGLIA, G.—STYAN, G. P. H.: *Rank conditions for generalized inverses of partitioned matrices*, Sankhyā Ser. A **86** (1974), 437–442.

[4] TIMM, N. H.: *The CGMANOVA model*, Comm. Statist. Theory Methods **26** (1997), 1083–1098.

[5] VERBYLA, A. P.—VENABLES, W. N.: *An extension of the growth curve model*, Biometrika **75** (1988), 129–138.

[6] VON ROSEN, D.: *Multivariate Linear Normal Models with Special Reference to the Growth Curve Model.* PhD Thesis, University of Stockholm, 1985.

[7] VON ROSEN, D.: *Maximum likelihood estimators in multivariate linear normal models*, J. Multivariate Anal. **31** (1989), 187–200.

[8] VON ROSEN, D.: *Moments for a multivariate linear model with an application to the growth curve model*, J. Multivariate Anal. **35** (1990), 243–259.

[9] VON ROSEN, D.: *Moments of maximum likelihood estimators in the growth curve model*, Statistics **22** (1991), 111–131.

[10] ŽEŽULA, I.: *Variance components estimability in the extended growth curve model*, Tatra Mt. Math. Publ. **26** (2003), 195–199.

Received October 3, 2006

Institute of Mathematics
Šafárik University
Jesenná 5
041 54 Košice
SLOVAKIA
E-mail: zezula@kosice.upjs.sk