RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **103** (1), 2009, pp. 161–166
Estadística e Investigación Operativa / Statistics and Operations Research

Some properties of the best linear unbiased estimators in multivariate growth curve models

Gabriela Beganu

Abstract The purpose of this article is to build a class of the best linear unbiased estimators (BLUE) of the linear parametric functions, to prove some necessary and sufficient conditions for their existence and to derive them from the corresponding normal equations, when a family of multivariate growth curve models is considered. It is shown that the classical BLUE known for this family of models is the element of a particular class of BLUE built in the proposed manner. The results are expressed in a convenient computational form by using the coordinate-free approach and the usual parametric representations.

Algunas propiedades de los estimadores lineales insesgados óptimos de los modelos con curva de crecimiento multivariantes

Resumen. El propósito del artículo es construir una clase de estimadores lineales insesgados óptimos (BLUE) de funciones paramétricas lineales para demostrar algunas condiciones necesarias y suficientes para su existencia y deducirlas de las correspondientes ecuaciones normales, cuando se considera una familia de modelos con curva de crecimiento multivariante. Se demuestra que la clase de los BLUE conocidos para esta familia de modelos es un elemento de una clase particular de los BLUE que se construyen de esta manera. Los resultados se presentan en un formato computacional adecuado usando un enfoque que es independiente de las coordenadas y las representaciones paramétricas usuales.

1 Introduction

Experimental techniques which consider the response of an individual over a period of time (or over different doses of some medicine) are generally named growth curve experiments. Their representation by growth curve models have been studied extensively in the literature because of their general aplicability (see [1, 6, 8, 10, 11]). The MANOVA models include the multivariate growth curve models but also the profile analysis models. The main difference is that in the profile analysis models the components of the vector of responses can be interchangeable whereas this question is not possible in the growth curve models.

The purpose of this article is to derive a class of the BLUE of linear parametric functions corresponding to a family of multivariate growth curve models. Some necessary and sufficient conditions for given estimable functions to be optimally estimable are proved and the BLUE of these functions are expressed using a coordinate-free approach.

The article is structurated as follows: In Section 2 a class of the BLUE of linear parametric functions is derived in a family of multivariate growth curve models. There are developed the properties of this set of the BLUE proving some necessary and sufficient conditions for their existence in this set.

Presentado por / Submitted by Francisco Javier Girón González-Torre.
Recibido / Received: 17 de noviembre de 2008. *Aceptado / Accepted*: 4 de marzo de 2009.
Palabras clave / Keywords: Best linear unbiased estimator, Linear parametric function.
Mathematics Subject Classifications: 62J05, 47A05.

The existence conditions for the optimal estimable parametric functions corresponding to this class of BLUE are given in Section 3 and the normal equations are also derived similarly for the general linear regression model. The results are applied to a particular choice of this class of the BLUE and it is shown that the maximum likelihood estimator (MLE)(which is the least squares estimator ([10, 3]) is a BLUE in this class. Therefore the conclusion is that the classical BLUE known for this family of linear models is the element of this particular class of BLUE built in the proposed manner.

2 A Class of Blue

We consider a family of multivariate growth curve models in the general form

$$E(Y) = ZBX' \tag{1}$$
$$\operatorname{cov}(Y) = Q \copyright \Sigma = V \tag{2}$$

where the between-individuals and the within-individual design matrices Z and X are known $n \times r$ and $p \times q$ matrices of full column rank, respectively, and B is an $r \times q$ matrix of unknown parameters. The rows of the observations Y are assumed to be independently and identically distributed with zero mean and the covariance matrix Σ. Q is an $n \times n$ symmetric and nonnegative definite (n.n.d.) matrix, which could arise in the context of some random effects.

In this paper there are used the following notations: $\mathcal{L}_{s,t}$ is the real vector space of all linear transformations on $\mathcal{S}$ to $\mathcal{T}$, where $\mathcal{S}$ and $\mathcal{T}$ stand for the s and t-dimensional real inner product spaces, respectively; $\mathcal{L}_{s,t}$ is endowed with the inner product $(A, E) = \operatorname{tr}(AE')$, where E' is the matrix of the adjoint operator E; the Kronecker product of the operators $A \in \mathcal{L}_{s,t}$ and $B \in \mathcal{L}_{u,v}$, such that $(A \copyright B)X = AXB'$ for all $X \in \mathcal{L}_{u,s}$, is a linear transformation on $\mathcal{L}_{u,s}$ to $\mathcal{L}_{v,t}$.

If we denote by

$$\mathcal{E} = \operatorname{sp}\{\, ZBX' \mid B \in \mathcal{L}_{q,r} \,\}$$

a linear subspace of a finite dimensional euclidian vector space $\mathcal{K} \subset \mathcal{L}_{p,n}$ and by

$$\mathcal{V} = \operatorname{sp}\{\, V : \mathcal{K} \longrightarrow \mathcal{K} \mid V \text{ a symmetric and n.n.d mapping} \,\}$$

a linear subspace of a set of all symmetric mappings from $\mathcal{K}$ to $\mathcal{K}$, then the model defined by the relations (1) and (2) can be expressed as an element of the set $M(\mathcal{E}, \mathcal{V})$ ([5]).

Corresponding to $Z \copyright X$, which is a linear operator on a finite dimensional Hilbert space $\mathcal{B} \subset \mathcal{L}_{q,r}$ to $\mathcal{K}$, the linear subspace $\mathcal{E}$ is the range of $Z \copyright X, \mathcal{E} = R(Z \copyright X)$.

Let V_0 be a maximal element of $\mathcal{V}$ (V_0 always exists in $\mathcal{V}$ [7]). Then $R(V) \subset R(V_0)$ for all $V \in \mathcal{V}$ and a new symmetric and n.n.d. operator W can be defined ([5]) such that

$$\mathcal{E} \subset R(W), \qquad R(V) \subset R(V_0) \subset R(W) \tag{3}$$

for all $V \in \mathcal{V}$. For the linear model (1), (2), operator W satisfying the properties (3) can be

$$W = V_0 + (ZZ') \copyright (XX'). \tag{4}$$

In the sequel it will be considered a parametric function (α, B), $\alpha, B \in \mathcal{B}$ that is optimally estimable [5], which means there exists an element $A \in \mathcal{K}$ such that (A, Y) is a BLUE of $E(A, Y)$ and a class of these BLUE is going to be built for the linear model $M(\mathcal{E}, \mathcal{V})$ defined by the relations (1) and (2).

Proposition 1 *(A, Y) is a* BLUE *of $E(A, Y)$ if and only if there exists an element $A_1 \in R(W)$ such that (A_1, Y) is a* BLUE *of $E(A, Y)$.*

PROOF. Let (A,Y) be a BLUE of $E(A,Y)$ with $A \in \mathcal{K}$. Then there exist $A_1 \in R(W)$ and $A_2 \in N(W)$ (the null space of the operator W), such that $A = A_1 + A_2$.

Since W satisfies the relations (3), we obtain from Theorem Farkas-Minkowski ([5]) that $N(W) \subset \mathcal{E}^{\perp}$ (the orthogonal complement of $\mathcal{E}$) and $N(W) \subset N(V)$ for all $V \in \mathcal{V}$.

Using these relations we can write that

$$E(A,Y) = (A_1, (Z \copyright X)B) = E(A_1, Y) \tag{5}$$

and

$$\begin{aligned} \operatorname{cov}\big((A,Y),(A,Y)\big) &= (A, VA) \\ &= (A_1 + A_2, VA_1) = (A_1, VA_1) \\ &= \operatorname{cov}\big((A_1,Y)(A_1,Y)\big) \end{aligned} \tag{6}$$

since $WA_2 = 0$ implies $VA_2 = 0$ for all $V \in \mathcal{V}$ and $(A_2, VA_1) = 0$ for $A_2 \in \mathcal{E}^{\perp}$ if and only if $VA_1 \in \mathcal{E}$ for all $V \in \mathcal{V}$. The last statement is the necessary and sufficient condition given by Theorem Lehman-Scheffé [9] for (A_1, Y) to be the BLUE of $E(A_1, Y)$.

From the relations (5) and (6) it follows that (A_1, Y) with $A_1 \in R(W)$ is a BLUE of $E(A_1, Y) = E(A,Y)$, $A \in \mathcal{K}$. ■

Let $\mathcal{C}$ be the class of the BLUE of $E(A,Y)$ for all $A \in R(W)$.

Some properties of $\mathcal{C}$ will be derived in the followings.

Proposition 2 *$(A,Y) \in \mathcal{C}$ if and only if*

$$A \in R[W^{+}(Z \copyright X)] \tag{7}$$

where W^{+} is the Moore-Penrose inverse of W.

PROOF. Let $A \in R(W)$. Then (A,Y) is a BLUE of $E(A,Y)$ accordingly to Proposition 1. By Theorem Lehmann-Scheffé the existence of the BLUE (A,Y) is equivalent to the condition $VA \in \mathcal{E}$ for all $V \in \mathcal{V}$. Since $R(ZZ' \copyright XX') = \mathcal{E}$, it follows, using the properties (3), that $WA \in \mathcal{E}$. This implies that $W^{+}WA \in W^{+}(\mathcal{E})$, which means that

$$A \in W^{+}(\mathcal{E}) = R[W^{+}(Z \copyright X)]$$

■

It is known ([2]) that

$$P = P_Z \copyright P_X = Z(Z'Z)^{-1}Z' \copyright X(X'X)^{-1}X'$$

and $M = I - P$ are the orthogonal projections onto $\mathcal{E}$ and $\mathcal{E}^{\perp}$, respectively (I is the identity matrix of corresponding orders).

Proposition 3 *Let $A \in R[W^{+}(Z \copyright X)]$ and let*

$$S(V) = (Z' \copyright X')W^{+}VMVW^{+}(Z \copyright X) \tag{8}$$

be a symmetric and n.n.d. mapping on $\mathcal{B}$ to $\mathcal{B}$ for all $V \in \mathcal{V}$. Then (A,Y) is a BLUE *of $E(A,Y)$ if and only if*

$$A \in R[W^{+}(Z \copyright X)(I - S^{-}(V)S(V))]$$

where $S^{-}(V)$ is a symmetric generalized inverse of $S(V)$ for all $V \in \mathcal{V}$.

PROOF. Let $A = W^+(Z \copyright X)B$ for some $B \in \mathcal{B}$. By Proposition 2 and Theorem Lehmann-Scheffé this means that $VA \in \mathcal{E}$ for all $V \in \mathcal{V}$. Then we have that

$$MVW^+(Z \copyright X)B = 0$$

and this relation is equivalent to $S(V)B = 0$ for all $V \in \mathcal{V}$, where $S(V)$ is given by (8).

Since $N[S'(V)] = R[S(V)]^\perp = R[I - S^-(V)S(V)]$ (by Theorem Farkas-Minkowski) we can write that $A \in R[W^+(Z \copyright X)(I - S^-(V)S(V))]$, $S(V)$ and $S^-(V)$ being symmetric mappings for all $V \in \mathcal{V}$. ■

Let $\{V_1, \ldots, V_m\}$ be a basis of $\mathcal{V}$.

Theorem 1 *$(A, Y) \in \mathcal{C}$ if and only if*

$$A \in R[W^+(Z \copyright X)(I - S^-S)] \tag{9}$$

where

$$S = \sum_{i=1}^{m} S(V_i) \tag{10}$$

PROOF. There are used Propositions 1, 2, 3 and a property of the null space of symmetric and n.n.d. mappings that $N(S) = \cap_{i=1}^m N[S(V_i)]$. ■

3 The Optimal Estimable Parametric Functions

Similarly to the results known for the general linear regression model, the elements $\alpha \in \mathcal{B}$ for which (α, B) is the optimal estimable parametric function will be derived and then $(\alpha, \hat{B})$ will be a BLUE of $E(A, Y)$ for all $A \in R(W)$ if $\hat{B}$ is a solution of the corresponding normal equations.

Proposition 4 *The linear parametric function (α, B) is optimally estimable for all $B \in \mathcal{B}$ if and only if*

$$\alpha \in R[(Z' \copyright X')W^-(Z \copyright X)(I - S^-S)]$$

where W^- is a symmetric generalized inverse of W.

PROOF. Accordingly to Theorem 1 $(A, Y) \in \mathcal{C}$ if and only if the relation (9) holds, which means that there exists a BLUE of the parametric function (α, B) if and only if

$$A = W^+(Z \copyright X)(I - S^-S)D \tag{11}$$

for some $D \in \mathcal{B}$ such that $E(A, Y) = (\alpha, B)$ for all $B \in \mathcal{B}$.

This can be written as

$$\begin{aligned} &W^+(Z \copyright X)(I - S^-S)D, (Z \copyright X)B) \\ &\qquad = ((Z' \copyright X')W^+(Z \copyright X)(I - S^-S)D, B) = (\alpha, B) \end{aligned}$$

for all $B \in \mathcal{B}$ if and only if

$$\alpha = (Z' \copyright X')W^+(Z \copyright X)(I - S^-S)D \tag{12}$$

for some $D \in B$.

Since $\mathcal{E} \subset R(W)$ we have that $Z \copyright X = WW^-(Z \copyright X)$, where W^- is a generalized inverse of W, that can be chosen to be a symmetric matrix. Then we have that

$$\begin{aligned} (Z' \copyright X')W^+(Z \copyright X) &= (Z' \copyright X')W^-WW^+WW^-(Z \copyright X) \\ &= (Z' \copyright X')W^-WW^-(Z \copyright X) \\ &= (Z' \copyright X')W^-(Z \copyright X) \end{aligned} \tag{13}$$

■

Theorem 2 *If $(A, Y) \in \mathcal{C}$ then $(\alpha, \hat{B})$ is a* BLUE *of $E(A, Y)$ if and only if $\hat{B}$ is a solution of the equation*

$$(I - S^{-}S)(Z' ©X')W^{-}(Z©X)B = (I - S^{-}S)(Z' ©X')W^{-}Y \tag{14}$$

PROOF. $(\alpha, \hat{B})$ is a BLUE of $E(A, Y)$, $A \in R(W)$, if and only if α can be expressed by the relation (12), which means that A verifies the equation (11) for some $D \in \mathcal{B}$. These statements manage to the relation

$$\begin{aligned} E(\alpha, \hat{B}) = (\alpha, B) &= \big((Z' ©X')W^{+}(Z©X)(I - S^{-}S)D, B\big) \\ &= \big((Z' ©X')A, B\big) \\ &= (A, (Z©X)B) \\ &= E(A, Y) \end{aligned}$$

or, equivalently, to the equality

$$\big(D, (I - S^{-}S)(Z' ©X')W^{+}(Z©X)\hat{B}\big) = \big(D, (I - S^{-}S)(Z' ©X')W^{+}Y\big)$$

for some $D \in \mathcal{B}$, which means that it is obtained the equation (14) if we allow for the relation (13) and the equality $(Z' ©X')W^{+}Y = (Z' ©X')W^{-}Y$. ■

Example 1 *It is proved in* [4] *that a* BLUE *of $E(Y)$ exists in the multivariate growth curve models defined by the relations* (1) *and* (2) *independently on the between-individuals matrix Z. This statement can justify the choice of the symmetric and n.n.d. operator defined by* (4) *as*

$$W = I_n ©I_p \tag{15}$$

Then the operator W verifies the properties (3).

In this case the operator S given by (10) *becomes*

$$S = \left(\sum_{i=1}^{m} Z'Q_iM_zQ_iZ\right) ©(X'X) = R©(X'X)$$

where $V_i = Q_i©I_p$, $i = 1, \ldots, m$ is a spanning set for $\mathcal{V}$, Q_i being a symmetric and n.n.d. matrix, $i = 1, \ldots, m$ and $M_Z = I_n - P_Z$.

Then the corresponding equation 14 *is*

$$\begin{aligned} &[(Z'Z)©(X'X) - (R^{-}RZ'Z)©(X'X)^{-}X'XX'X]B \\ &\qquad = [Z' ©X' - (R^{-}RZ')©(X'X)^{-}X'XX']Y \end{aligned}$$

and a solution is given by

$$\hat{B} = [(Z'Z)^{-}Z' ©(X'X)^{-}X']Y \tag{16}$$

which is the classical MLE *and it is the same as the least squares estimator of B.*

It can be noticed that $(\alpha, \hat{B})$, with $\hat{B}$ determined above (16)*, is an element of the class $\mathcal{C}$ derived for a certain choice* (15) *of the operator W.*

References

[1] ARNOLD, S. F., (1979). A coordinate-free approach to finding optimal procedures for repeated measures designs, *Ann. Statist.*, **7**, 812–822.

[2] BEGANU, G., (2006). A two-stage estimator of individual regression coefficients in multivariate linear growth curve models, *Rev. Acad. Colomb. Cienc.*, **30**, 548–554.

[3] BEGANU, G., (2007). On the equality of the ordinary least squares estimators and the best linear unbiased estimators in multivariate growth-curve models, *Rev. R. Acad. Cien. Serie A. Mat.*, **101**, 63–70.

[4] BEGANU, G., (2007). The best linear unbiased estimators of regression coefficients in a multivariate growth-curve model. A coordinate-free approach, *Rev. Acad. Colomb Cienc.*. **31**, 257–273.

[5] DRYGAS, H., (1975). Estimation and prediction for linear models in general spaces, *Math. Operationsforsch. Statist.*, **6**, 301–324.

[6] KURATA, H. AND KARIYA, T., (1996). Least upper bound for the covariance matrix of a generalized least squares estimator in regression with applications to a seemingly unrelated regression model and a heteroscedastic model, *Ann. Statist.*, **24**, 1547–1559.

[7] LAMOTTE, L. R., (1977). A canonical form for the general linear model, *Ann. Statist.*, **5**, 787–789.

[8] LANGE N. AND LAIRD N. M., (1989). The effect of covariance structure on variance estimation in balanced growth-curve models with random parameters, *J. Amer. Statist. Assoc.*, **84**, 241–247.

[9] LEHMANN E. AND SCHEFFÉ, H., (1950). Completeness, similar regions and unbiased estimation, *Sankhya*, **10**, 305–340.

[10] REINSEL, C. G., (1982). Multivariate repeated-measurement or growth curve models with multivariate random effects covariance structure, *J. Amer. Statist. Assoc.*, **77**, 190–195.

[11] SALA-I-MARTIN, X., DOPPELHOFER, G. AND MILLER, R. I., (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach, *American Econ. Rev*. **94**, 813–835.

Gabriela Beganu
Academy of Economic Studies
Department of Mathematics
Piaţa Romană, nr. 6
Bucharest, Romania
gabriela_beganu@yahoo.com