RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **99** (2), 2005, pp. 187–193
Estadística e Investigación Operativa / Statistics and Operations Research

A Gram-Schmidt Orthogonalizing Process of Design Matrices in Linear Models as an Estimating Procedure of Covariance Components

Gabriela Beganu

Abstract. It is considered a multivariate balanced mixed linear model without interaction for which the matrices of the quadratic forms required to estimate the covariance components are expressed by linear operators on finite dimensional inner product spaces.

The purpose of this article is to prove that the quadratic forms obtained by a Gram - Schmidt orthogonalizing process of design matrices are linear combinations of the quadratic forms derived by the generalized fitting constants method. Some sufficient conditions for the existence of non-negative best quadratic unbiased estimators (BQUE) for linear functions of covariance components are derived in a coordinate-free approach.

Un proceso de ortogonalización de Gram-Schmidt de diseño de matrices en modelos lineales como procedimiento para estimar la covarianza de las componentes

Resumen. Se considera un modelo lineal mixto multivariante equilibrado sin interacción para el que las matrices de las formas cuadráticas necesarias para estimar la covarianza de las componentes se expresan mdiante operadores lineales en espacios con producto interior de dimensión finita.

El propósito de este artículo es demostrar que las formas cuadráticas obtenidas por el proceso de ortogonalización de Gram-Schmidt de las matrices de diseño son combinaciones lineales de las formas cuadráticas derivadas del método generalizado de ajuste de constantes. Se deducen algunas condiciones suficientes para la existencia de mejores estimadores cuadrticos no sesgados (BQUE) para funciones lineales de componentes de covarianza utilizando un método libre-coordenadas.

1 Introduction

The estimating procedures of covariance matrices in linear models may encounter the problem of negative definite quadratic estimators.

Rao and Kleffe [13] developed estimation methods to obtain minimum norm quadratic unbiased estimators (MINQUE), minimum variance quadratic unbiased estimators (MIVQUE), maximum likelihood estimators (MLE), but these are not necessarely positive definite quadratic forms.

A necessary and sufficient condition for admissible estimators of variance components to be non-negative was established by Klonecki and Zontek [11] in univariate unbalanced mixed models. Using models with only two covariance components matrices, Amemiya [1] and Andesrson, Anderson and Olkin [2]

Presentado por F. J. Girón.
Recibido: 19 de octubre de 2005. *Aceptado*: 16 de noviembre de 2005.
Palabras clave / Keywords: Orthogonal projection, orthogonal complement, best quadratic unbiased estimator
Mathematics Subject Classifications: Primary 62J10. Secondary 47A05.
© 2005 Real Academia de Ciencias, España.

derived conditions for non-negativity of MLE of covariance components. For any balanced mixed linear model, Calvin and Dykstra [8] proposed an estimating procedure of covariance matrices subject to the restriction that the difference between certain pairs of matrices are non-negative definite. Wu and Pourahmadi [21] derived nonparametric positive definite estimators of covariance matrices using autoregressive models of a suitable order and assuming the stationarity of processes. The MLE of covariance components in a completely balanced multivariate multi-way random effect model without interaction were obtained by Tsai [20] using a new parametrization for covariance matrices.

In this paper it is considered a multivariate mixed linear model without interaction. The unbiased estimators of covariance matrices obtained by a generalization of Henderson method III are presented in Section 2. There are also proved some results regarding the orthogonal projections used to express these estimators.

Section 3 deals with determining the linear operators of the quadratic forms founded by Tan [19] by an orthogonalizing process of the design matrices corresponding to the model. It is shown that the two estimating procedures - the generalized fitting constants method and the orthogonalizing process - have the same solution (when it exists). The sufficient conditions for the existence of the non-negative BQUE of linear parametric functions are expressed in Section 4 using the orthogonal projections defined for the generalized fitting constants method. The results are illustrated on a two-factor random effects model without interaction - univariate and multivariate cases - and some concluding remarks are made in Section 5.

2 The generalized fitting constants method

Let

$$Y = X\beta_0 + \sum_{h=1}^{k} Z_h\beta_h + e \tag{1}$$

be a multivariate mixed linear model, where X and Z_h are $N \times m$ and $N \times n_h$ design matrices, respectively, β_0 is an $m \times p$ matrix of unknown parameters, β_h an $n_h \times p$ matrix of random variables for $h = 1, \dots k$ and e is an $N \times p$ matrix of erorrs. It is assumed that the rows of β_h and e are independent and identically normal distributed random vectors with zero means and corresponding non-singular covariance matrices Σ_h, $h = 1, \dots k$ and $\Sigma_e = \Sigma_{k+1}$, respectively. Then the random matrix Y has the expected value

$$E(Y) = X\beta_0 \tag{2}$$

and the covariance matrix

$$\operatorname{cov}(Y) = \sum_{h=1}^{k+1} (Z_h Z_h') \otimes \Sigma_h \tag{3}$$

where it is considered that $Z_{k+1}Z_{k+1}' = I$ is the identity $N \times N$ matrix and "$\otimes$" is the Kronecker matrix product.

Concerning vector space notions which are utilized in the sequel, we mention a few at this point (Halmos [9]).

Let $\mathcal{L}_{p_1,p_2}$ be the finite dimensional linear space of all $p_2 \times p_1$ real matrices which is endowed with the inner product $\langle A, B\rangle = \operatorname{tr}(AB')$ for arbitrary $A, B \in \mathcal{L}_{p_1,p_2}$ and let P be a linear operator from $\mathcal{L}_{p_1,p_2}$ to L_{q_1,q_2}. The adjoint operator of P is the linear operator P^* from $\mathcal{L}_{q_1,q_2}$ to $\mathcal{L}_{p_1,p_2}$ having the property $\langle P^*A, B\rangle_1 = \langle A, PB\rangle_2$ for all $A \in \mathcal{L}_{q_1,q_2}$ and $B \in \mathcal{L}_{p_1,p_2}$. The inner products $\langle\cdot,\cdot\rangle_1$ $\langle\cdot,\cdot\rangle_2$ are defined on $\mathcal{L}_{p_1,p_2}$ and L_{q_1,q_2}, respectively. In the sequel it will be used the same trace inner product for all linear spaces. The range of the linear operator P is the linear subspace $R(P)$ of $\mathcal{L}_{q_1,q_2}$ spanned by the columnes of P and the rank of P is denoted by $r(P)$.

The orthogonal complement of a non-empty subset A with respect to a certain inner product is denoted by $A^{\perp}$.

In order to estimate the covariance components $\Sigma_1, \dots, \Sigma_{k+1}$ by applying the Henderson method III (Henderson [10], Searle [15]) it will be used the generalized least squares estimation procedure for every i submodel constructed from model (1) with the design matrix $U_i = (X, Z_1, \dots, Z_i)$ and having $\Theta_i = (\beta_0', \beta_1', \dots, \beta_i')$ an $p \times (m + \sum_{h=1}^{i} n_h)$ matrix of unknown parameters, $i = 1, \dots, k$. Then the random matrix Y in the i submodel (considered as a fixed linear model) has expectation

$$E(Y) = X\beta_0 + \sum_{h=1}^{i} Z_h\beta_h = U_i\Theta_i \tag{4}$$

and covariance matrix

$$\text{cov}(Y) = I \otimes \Sigma_e \tag{5}$$

for all $i = 1, \dots, k$. If we denote $U_0 = X$ and $\Theta_0 = \beta_0$, then there are $k + 1$ submodels of model (1). For the i submodel $\hat{\Theta}_i = (U_i'U_i)^{-}U_i'Y$ is the ordinary least squares estimator of Θ_i, where $(U_i'U_i)^{-}$ is a g-inverse of $U_i'U_i$ if U_i is not of full column rank, $i = 0, 1, \dots, k$. Then the linear operator

$$P_i = U_i(U_i'U_i)^{-}U_i' \tag{6}$$

from $\mathcal{L}_{p,N}$ to $\mathcal{L}_{p,N}$ is an orthogonal projection on $R(U_i)$ for all $i = 0, 1, \dots, k$.

Lemma 1 *If the linear operator P_i is given by the relation* (6)*, then*

$$P_i = P_{i-1} + (I - P_{i-1})Z_iT_i^{-}Z_i'(I - P_{i-1}) \tag{7}$$

where

$$T_i = Z_i'(I - P_{i-1})Z_i \tag{8}$$

for all $i = 1, \dots, k$.

PROOF. The recurrence formula (7) is proved using the formula for obtaining a generalized inverse of a partitioned symmetric matrix [14] and noticeing that $U_i = (U_{i-1}, Z_i)$ for $i = 1, \dots, k$. ■

Corollary 1 *If the linear operator P_i given by* (6) *is an orthogonal projection on $R(U_i)$, then it is an orthogonal projection on $R(U_h)$ for $h = 0, 1, \dots, i-1$ and $i = 1, \dots, k$.*

PROOF. For $i = 1$ we have $P_1U_1\Theta_1 = U_1\Theta_1$. Then $P_1X\beta_0 = X\beta_0$ because $P_0 = X(X'X)^{-}X'$ is an orthogonal projection on $R(U_0)$.

It will be easily proved that $P_iU_h\Theta_h = U_h\Theta_h$ for all $h = 0, 1, \dots, i-1$ using the relation (7). ■

Lemma 2 *If the linear operator P_i is given by the relation* (6)*, then $P_i - P_{i-1}$ is an orthogonal projection on $R(U_{i-1})^{\perp}$ for all $i = 1, \dots, k$.*

PROOF. It is used Corollary (1). ■

Corollary 2 *If the linear operator P_i is defined by* (6)*, then $P_i - P_{i-1}$ is an orthogonal projection on*

$$R(X)^{\perp} \bigcap \left[\bigcap_{h=1}^{i-1} R(Z_h)^{\perp}\right] \qquad \textit{for } i = 1, \dots, k.$$

Corollary 3 *If the linear operator P_i is defined by* (6)*, then $P_i - P_{i-1}$ is an orthogonal projection on*

$$R(XX^*)^{\perp} \bigcap \left[h = 1^{i-1} R(Z_hZ_h^*)^{\perp}\right] \qquad \textit{for } i = 1, \dots, k.$$

Applying the fitting constants method of Henderson in some univariate mixed linear models, Seely [16, 17] and Seely and Zyskind [18] provided necessary and sufficient conditions for the existence of the quadratic unbiased estimators of variance-covariance components. The results could be extended to the multivariate case of the model (1) under the assumptions (2), (3) (Beganu [3, 4, 5, 6]). It was proved that the quadratic unbiased estimators of covariance components obtained by the generalized fitting constant method are the solutions of the system

$$\begin{cases} \sum_{h=i}^{k} \mathrm{tr}[Z_h'(P_k - P_{i-1})Z_h] \cdot \Sigma_h + [r(U_k) - r(U_i)] \cdot \Sigma_e = Y'(P_k - P_{i-1})Y, \qquad i = 1, \ldots, k \\ [N - r(U_k)] \cdot \Sigma_e = Y'(I - P_k)Y \end{cases} \tag{9}$$

where it was used a result obtained by Neudecker [12].

Thus the system can be solved when it is consistent (under certain conditions [7]) but the solution has not necessarely non-negative definite components [17].

3 Orthogonalizing process-an estimating procedure

The purpose of this section is to prove that the quadratic estimators of covariance matrices obtained by Tan [19] and by Henderson method III coincide for model (1).

It can be shown that the symmetric matrices founded by a Gram-Schmidt iterative method to orthogonalize the design matrices of model (1) verify the relations

$$\begin{aligned} W_i = \ & \prod_{h=1}^{i-1}(I - W_h)(I - P_0)Z_i[Z_i'(I - P_0)\prod_{h=1}^{i-2}(I - W_h)(I - W_{i-1}) \cdot \\ & \prod_{h=1}^{i-2}(I - W_h)(I - P_0)Z_i]^{-} \cdot Z_i'(I - P_0)\prod_{h=1}^{i-1}(I - W_h) \end{aligned} \tag{10}$$

and hence it can be written

$$P_i = P_0 + \sum_{h=1}^{i} W_h \tag{11}$$

where P_i is given by (6) for all $i = 1, \ldots, k$.

Theorem 1 *If P_i is the orthogonal projection* (6) *on $R(U_i)$ and W_i is the linear operator from $L_{p,N}$ to $L_{p,N}$ verifying the relation* (10)*, then*

$$W_i = (I - P_{i-1})Z_i T_i^{-} Z_i'(I - P_{i-1}) = P_i - P_{i-1} \tag{12}$$

where T_i is given by (8) *for $i = 1, \ldots, k$.*

PROOF. It is easy to prove that

$$\prod_{h=1}^{i-1}(I - W_h)(I - P_0) = I - P_{i-1}$$

if the results of Corollary 1 and relation (11) are used for $i = 1, \ldots, k$. Then

$$\begin{aligned} & Z_i'(I - P_0)\prod_{h=1}^{i-1}(I - W_h)\prod_{h=1}^{i-2}(I - W_h)(I - P_0)Z_i \\ & \qquad = Z_i'(I - P_{i-1})(I - P_{i-2})Z_i = Z_i'(I - P_{i-1})Z_i = T_i \end{aligned}$$

for all $i = 1, \ldots, k$. Hence the first equality from (12) is true and the second equality results from (11).

The generalized quadratic forms corresponding to the symmetric matrices W_i given by (12) will have the expected values

$$E(Y'W_iY) = \sum_{h=i}^{k} tr(Z_h'W_iZ_h) \cdot \Sigma_h + [r(U_i) - r(U_{i-1}] \cdot \Sigma_e$$

for $i = 1, \cdots k$ and

$$E[Y'(I - W_k)Y] = [N - r(U_k)] \cdot \Sigma_e.$$

These expressions could be obtained from the assumptions (2) and (3) and the results of Corollaries 2 and 3. Then the estimating equations founded by Gram-Schmidt method to orthogonalize the design matrices of model (1) become

$$\begin{cases} \sum_{h=i}^{k} \mathrm{tr}(Z_h'W_iZ_h) \cdot \Sigma_h + [r(U_i) - r(U_{i-1})] \cdot \Sigma_e = Y'W_iY \qquad i = 1, \ldots, k \\ [N - r(U_k)] \cdot \Sigma_e = Y'(I - W_k)Y \end{cases} \tag{13}$$

It is easy to see using the relation (12) that the estimating equations (9) and (13) yield the same unbiased quadratic estimators of $\Sigma_1, \ldots, \Sigma_k, \Sigma_e$.

4 Non-negative BQUE of covariance components

The existence of non-negative BQUE of covariance components is not generally considered for Henderson method III estimating procedure. For particular model (1) Tan [19] proved in Theorem 3.1 some sufficient conditions to exist non-negative BQUE for linear combinations of covariance components. These conditions can be enounced in the framework of coordinate-free approach as follows:

Theorem 2 *Let W_i be the linear operator* (10) *for $i = 1, \ldots, k$. If:*

(i) W_iD is an orthogonal projection on $\bigcap_{h=i}^{k+1} R(Z_hZ_h^*)^{\perp}$, $i = 1, \ldots, k$, for any linear operator D which is an orthogonal projection on

$$R(XX^*)^{\perp} \bigcap \left[\bigcap_{h=1}^{k+1} R(Z_hZ_h^*)^{\perp} \right],$$

(ii) the matrices W_i and $Z_h'W_iZ_h$ have equal diagonal elements for all $h = i, \ldots, k$,

then $Y'W_iY$ is a non-negative BQUE of $E(Y'W_iY)$ for $i = 1, \ldots, k$.

Corollary 4 *If the conditions (i) and (ii) hold, then $Y'(P_k - P_{i-1})Y$ is a non-negative BQUE for the linear function of $\Sigma_i, \ldots, \Sigma_k$, and Σ_e expressed by the left-side of the relations* (9) *for all $i = 1, \ldots, k$.*

Corollary 5 *If P_k has equal diagonal elements then $Y'(I - P_k)Y$ is a non-negative BQUE of $[N - r(U_k)] \cdot \Sigma_e$.*

These results can be extended for linear parametric functions $\sum_{i=1}^{k+1} \lambda_i \Sigma_i$, where $\lambda = (\lambda_1, \ldots, \lambda_{k+1})' \in \mathbb{R}^{k+1}$ and denoting the symmetric matrices of the quadratic forms from (9) by $Q_i = P_k - P_{i-1}$ for $i = 1, \ldots, k$ and $Q_{k+1} = I - P_k$.

Theorem 3 *If the linear operator W_i verifies conditions (i) and (ii) and if there exists $\rho \in \mathbb{R}^{k+1}$ such that the equations*

$$\begin{cases} \sum_{h=1}^{i} tr(Z_i' Q_h Z_i) \cdot \rho_h = \lambda_i, \qquad i = 1, \ldots, k \\ \sum_{h=1}^{k} [r(U_k) - r(U_{h-1})] \cdot \rho_h + [N - r(U_k)] \cdot \rho_{k+1} = \lambda_{k+1} \end{cases} \tag{14}$$

are satisfied for every $\lambda \in \mathbb{R}^{k+1}$, then the linear function $\sum_{i=1}^{k+1} \rho_i Y' Q_i Y$ is a non-negative BQUE of $\sum_{i=1}^{k+1} \lambda_i \Sigma_i$.

PROOF. The results of Theorem (2) in [17] and of Corollary (2) in [7] are used in order to obtain that a linear parametric function $\sum_{i=1}^{k+1} \lambda_i \Sigma_i$ is estimable if the equations (14) are consistent for every $\lambda \in \mathbb{R}^{k+1}$. ■

References

[1] Amemiya, Y. (1985). What should be done when an estimated between-grup covariance matrix is not nonnegative definite?, *Amer. Statistic.*, **39**, 112–117.

[2] Anderson, B. M., Anderson, T. W. and Olkin, I. (1986). Maximum likelihood estimators and likelihood ratio criteria in multivariate components of variance, *Ann. Statist.*, **14**, 405–417.

[3] Beganu, G. (1987). Estimation of regression parameters in a covariance linear model, *Stud. Cerc. Mat.*, **39** , 3–10.

[4] Beganu, G. (1987). Estimation of covariance components in linear models. A coordinate–free approach, *Stud. Cerc. Mat.*, **39**, 228–233.

[5] Beganu, G. (1992). A model of multivariate analysis of variance with applications to medicine, *Econom. Comput. Econom. Cybernet. Stud. Res.*, **7** , 35–40.

[6] Beganu, G. (2003). The existence conditions of the best linear unbiased estimators of the fixed factor effects, *Econom. Comput. Econom. Cybernet, Stud. Res.*, **36**, 95–102.

[7] Beganu, G. Quadratic estimators of covariance components in a multivariate mixed linear model, *Statist. Meth. App.* (to appear).

[8] Calvin, J. A. and Dykstra, R. L., (1991). Maximum likelihood estimation of a set of covariance matrices under Löwner order restrictions with applications to balanced multivariate variance components models, *Ann. Statist.*, **19**, 850–869.

[9] Halmos, P. R., (1958). *Finite Dimensional Vector Spaces*, 2nd ed., Van Nostrandan, Princeton.

[10] Henderson, C. R., (1953). Estimation of variance and covariance components, *Biometrics*, **9**, 226–252.

[11] Klonecki, W. and Zontek, S. (1992). Admissible estimators of variance components obtained via submodels, *Ann. Statist.*, **20**, 1454–1467.

[12] Neudecker, H., (1990). The variance matrix of a matrix quadratic form under normality assumptions. A derivation based on its moment–generating function, *Math. Operationsforsch. Statist., ser. Statistics*, **3**, 455–459.

[13] Rao, C. R. and Kleffe, J., (1988). *Estimation of Variance Components and Applications*, North–Holland, Amsterdam.

[14] Searle, S. R. (1971). *Linear Models*, Willey, New York.

[15] Searle, S. R. (1971). Topics in variance components estimation, *Biometrics*, **27**, 1–76.

[16] Seely, J., (1970). Linear spaces and unbiased estimation, *Ann. Math. Statist.*, **41**, 1725–1734.

[17] Seely, J., (1970). Linear spaces and unbiased estimation. Application to the mixed linear model, *Ann. Math. Statist.*, **41**, 1735–1748.

[18] Seely, J. and Zyskind, G., (1971). Linear spaces and minimum variance unbiased estimation, *Ann. Math. Statist.*, **42**, 691–703.

[19] Tan, W. Y., (1979). On the quadratic estimation of covariance matrices in multivariante linear models, *J. Multiv. Anal.*, **9**, 452–459.

[20] Tsai, M. T., (2004). Maximum likelihood estimation of covariance matrices under simple tree ordering, *J. Multiv. Anal.*, **89**, 292–301.

[21] Wu, W. B. and Pourahmadi, M., (2003). Nonparametric estimation of large covariances matrices of longitudinal data, *Biometrika*, **90**, 831–844.

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