RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **101** (2), 2007, pp. 159–166
Estadística e Investigación Operativa / Statistics and Operations Research

On some alternative forms equivalent to Kruskal's condition for OLSE to be BLUE

Gabriela Beganu

Abstract. The necessary and sufficient condition for the ordinary least squares estimators (OLSE) to be the best linear unbiased estimators (BLUE) of the expected mean in the general univariate linear regression model was given by Kruskal (1968) using a coordinate-free approach. The purpose of this article is to present in the same manner some alternative forms of this condition and to prove two of the Haberman's equivalent conditions in a different and simpler way. The results obtained in the general univariate linear regression model are applied to a family of multivariate growth curve models for which the problem of the equality between OLSE and BLUE is treated in a coordinate-free approach.

Sobre algunas formas alternativas equivalentes a la condición de Kruskal para los OLSE y los BLUE

Resumen. La condición necesaria y suficiente para que los estimadores ordinarios de mínimos cuadrados (OLSE) sean los mejores estimadores lineales insesgados (BLUE) de la media esperada en el modelo general de regresión lineal univariante la dio Kruskal (1968) usando un enfoque que no dependía de las coordenadas. El propósito de este artículo es presentar, del mismo modo, algunas formas alternativas de esta condición y demostrar dos de las condiciones equivalentes de Haberman de un modo diferente pero más sencillo. Los resultados que se obtienen para el modelo general de regresión lineal univariante se aplican a una familia de modelos de crecimiento multivariante para los que el problema de la igualdad entre los OLSE y los BLUE se trata usando un enfoque que no depende de las coordenadas.

1 Introduction

The problem of the equality between OLSE and BLUE (or Gauss-Markov estimators) of the expected mean in linear regression models is treated in a coordinate-free approach.

Significant contributions in solving this problem were brought by Haberman [12], Arnold [1], Baksalary and van Eijnsbergen [3], Puntanen and Styan [16], Puntanen, Styan and Tian [17] among many others, but Zyskind [23] proved the first general theorem for OLSE to be BLUE when the covariance matrix is non-negative definite and the design matrix of the model does not necessarily have full column rank. Zyskind and Martin [24] and Seely [21, 22] obtained significant general results on unbiased estimation restricted to finite dimensional linear spaces.

One of the initial papers in which geometric aspects of the Gauss-Markov estimation were used in the general univariate linear model is Kruskal's [14]. In the same manner Eaton [11], Drygas [9, 10], Baksalary and Kala [2], Beganu [4, 5, 6] gave some alternative forms, as well as their extensions, of the necessary and

Presentado por David Ríos Insua.
Recibido: 29 de enero de 2007. *Aceptado*: 7 de marzo de 2007.
Palabras clave / Keywords: Ordinary least squares estimator, best linear unbiased estimator, orthogonal projection.
Mathematics Subject Classifications: 62J05, 47A05

sufficient conditions for BLUE to exist and to be identical to OLSE of the expected mean in multivariate linear regression models.

The purpose of this article is to present some alternative forms of Kruskal's condition using a coordinate-free approach and to apply them to a family of multivariate linear regression models.

Classical results for the general univariate linear regression model are outlined in Section 2 and another proof of some alternative forms given by Haberman [12] for OLSE to be BLUE are obtained. It is also proved an equivalent condition to Haberman's conditions.

In Section 3 the Kronecker matrix product and the Kronecker product of linear transformations are applied to describe the covariance structure of a random matrix of observations in a family of multivariate growth curve models and to construct the corresponding orthogonal projections. It is used the necessary and sufficient condition for OLSE to be identical to BLUE obtained by Beganu (to appear [8]) for this model and it will be proved that the equivalent conditions in Section 2 can be applied to this family of multivariate linear regression models. It will be also shown that these necessary and sufficient conditions for the equality of OLSE and BLUE do not depend on the between-individuals design matrix of the multivariate growth-curve models.

2 Equivalent conditions for OLSE to be BLUE

The general univariate linear regression model assumes that the random vector Y of observations is valued in an n-dimensional real vector space W endowed with the inner product $(\cdot\,,\cdot)$. The expected mean μ of Y lies in a linear manifold Ω of W and the covariance operator of Y is $\Sigma = \sigma^2 V$, where V is a known positive definite linear transformation on W to W and σ^2 is a known or unknown positive number.

Using a coordinate-free approach, the OLSE and the BLUE of μ are defined by the orthogonal projections on Ω relative to the inner products denoted by (x,z) and $((x,z)) = (x, V^{-1}z)$ for any $x, z \in W$. The orthogonal projections P and Q on Ω relative to the two inner products, respectively, are the linear transformations on W to Ω such that Px and Qx are the unique elements in Ω which satisfy the relations $(x - Px, z) = 0$; and $((x - Qx, z)) = 0$ for all $x \in W, z \in \Omega$. Then the OLSE of μ is $\hat{\mu}_{\text{OLSE}} = PY$ and the BLUE of μ is $\hat{\mu}_{\text{BLUE}} = QY$.

It is known (Zyskind [23], Kruskal [14]) that

$$\hat{\mu}_{\text{OLSE}} = \hat{\mu}_{\text{BLUE}} \tag{1}$$

if and only if $V\Omega \subset \Omega$ with alternatives (Haberman [12])

$$V^{-1}\Omega \subset \Omega \tag{2}$$

and $V\Omega^{\perp} \subset \Omega^{\perp}$, $V^{-1}\Omega^{\perp} \subset \Omega^{\perp}$ by non-singularity of V, where $\Omega^{\perp}$ is the orthogonal complement of Ω.

Haberman [12] also proved some other alternative forms of the necessary and sufficient conditions for (1) in the univariate case of the general linear regression model.

One of these alternative forms can be proved in another manner using the definitions and the properties of the orthogonal projections P and Q on Ω.

Proposition 1 *The equality* (1) *holds if and only if*

$$(x, PV^{-1}Pz) = (x, PV^{-1}z) \tag{3}$$

for all $x, z \in W$.

PROOF. Q is the orthogonal projection on Ω, which means that $((Px, Qz)) = ((Px, z))$ for all $x, z \in W$. Then the relation

$$((Px, Qz)) = ((Px, Pz)) + ((Px, (I - P)z))$$

can be obtained and it is equivalent to

$$((Px, Qz)) = ((Px, Pz)) \tag{4}$$

for all $x, z \in W$ if and only if $((Px, (I-P)z)) = 0$, namely

$$((Px, Pz)) = ((Px, z)), \qquad x, z \in W \tag{5}$$

which is identical to condition (3).

But, P being symmetric, the two terms in (4) are

$$((Px, Qz)) = (x, PV^{-1}Qz) = (x, V^{-1}Qz) = ((x, Qz))$$

and

$$((Px, Pz)) = (x, PV^{-1}Pz) = (x, V^{-1}Pz) = ((x, Pz))$$

for all $x, z \in W$ if and anly if the relation (2) is accomplished.

Therefore the equality (4) becomes

$$((x, \hat{\mu}_{\text{BLUE}})) = ((x, \hat{\mu}_{\text{OLSE}}))$$

for all $x \in W$ and $z = y \in W$. ■

Condition (3) represents the alternative form (5) obtained by Haberman [12] in Theorem 2.

An alternative form of condition (3) is the following:

Proposition 2 *The estimators $\hat{\mu}_{\text{BLUE}}$ and $\hat{\mu}_{\text{OLSE}}$ are identical if and only if*

$$(x, V^{-1}Q^*z) = (x, PV^{-1}Q^*z) \tag{6}$$

for all $x, z \in W$.

PROOF. It is known that P and Q satisfy the relation $PQ = Q$. Then, using the definition of the adjoint operator of Q, we have that

$$((Qx, Pz)) = ((x, Q^*Pz)) = ((x, Q^*z)) = ((Qx, z))$$

for all $x, z \in W$. But Q is the orthogonal projection on Ω which means that $((Qx, Pz)) = ((x, Pz))$ for all $x, z \in \Omega$. It follows that the equality

$$((Qx, z)) = ((x, Pz))$$

can be obtained or, an equivalent equality,

$$((Qx, z)) = ((Px, Pz)) + (((I-P)x, Pz)) \tag{7}$$

for all $x, z \in W$.

The second term of the right-hand side in (7) is null if and only if $V^{-1}\Omega \subset \Omega$ and can be written as

$$\begin{aligned} 0 &= (((I-P)x, Pz)) = (((I-P)x, Q^*Pz)) \\ &= ((I-P)x, V^{-1}Q^*z) = (x, (I-P)V^{-1}Q^*z) \end{aligned}$$

which is equivalent to condition (6), or to another form of (6)

$$((Qx, Pz)) = ((Px, Pz)), \qquad x, z \in W. \tag{8}$$

Then the equality (7) becomes

$$((Qx, z)) = ((Px, Pz))$$

for all $x, z \in W$ if and only if $V^{-1}\Omega \subset \Omega$. But, according to Proposition 1, $V^{-1}\Omega \subset \Omega$ holds iff the relation (5) is accomplished. It follows that

$$((\hat{\mu}_{\text{BLUE}}, z)) = ((\hat{\mu}_{\text{OLSE}}, z))$$

for all $z \in W$ and $x = y \in W$. ■

Proposition 3 *The necessary and sufficient condition* (6) *for* $\hat{\mu}_{\text{OLSE}}$ *to be* $\hat{\mu}_{\text{BLUE}}$ *is equivalent to*

$$(x, PV^{-1}Pz) = (x, V^{-1}Pz) \tag{9}$$

for all $x, z \in W$.

PROOF. Using the definition of the inner product $((\cdot\,,\cdot))$ in W, the second term in the right-hand side in (7) can be written as

$$0 = (((I-P)x, Pz)) = ((I-P)x, V^{-1}Pz) = (x, (I-P)V^{-1}Pz)$$

or

$$0 = ((x, Pz)) - ((Px, Pz)) \tag{10}$$

for all $x, z \in W$. ■

Condition (9) was obtained by Haberman [12] (condition (6) in Theorem 2) using the symmetry of the linear operator $PV^{-1}(I-P)$.

3 Equivalent conditions in multivariate growth curve models

In the sequel the necessary and sufficient conditions for the equality between $\hat{\mu}_{\text{OLSE}}$ and $\hat{\mu}_{\text{BLUE}}$ expressed in Propositions 1, 2 and 3 will be verified using an example of multivariate growth curve model and some algebrical preliminaries (Eaton [11], Halmos [13], Rao [18]) are needed for this purpose.

Let $\mathcal{L}_{p_1,p_2}$ be the vector space of $p_2 \times p_1$ real matrices endowed with the inner product $< A, B >= \text{tr}(AB')$ for all $A, B \in \mathcal{L}_{p_1,p_2}$. (The same trace inner product will be considered for all real vector spaces $\mathcal{L}_{p_1,p_2}$). The Kronecker matrix product is defined as usual: if $A \in \mathcal{L}_{p_1,p_2}$ and $B \in \mathcal{L}_{q_1,q_2}$, then $A \otimes B = (a_{i,j}B) \in \mathcal{L}_{p_1q_1,p_2,q_2}$.

If V_i is a p_i-dimensional real vector space with the inner product $(\cdot\,,\cdot)_i$, $i = 1, 2$, then $\mathcal{L}_{p_1,p_2}$ is also the real vector space of the linear transformations on V_1 to V_2. The Kronecker product of the linear operators $A \in \mathcal{L}_{p_2,p_2}$ and $B \in \mathcal{L}_{p_1,p_1}$ is the linear operator on $\mathcal{L}_{p_1,p_2}$ to $\mathcal{L}_{p_1,p_2}$ such that $(A \copyright B)T = ATB^*$ for every $T \in \mathcal{L}_{p_1,p_2}$, where B^* is the adjoint operator of B relative to the trace inner product in $\mathcal{L}_{p_1,p_1}$. The composition of two Kronecker operators products is

$$(A_1 \copyright B_1) \circ (A_2 \copyright B_2) = (A_1A_2) \copyright (B_1B_2)$$

and the adjoint operator is

$$(A \copyright B)^* = A^* \copyright B^*.$$

It has also to be reminded the property of any $A \in \mathcal{L}_{p_1,p_2}$ which has rank one to be expressed $A = a_2 \otimes a_1$ with $a_1 \in V_1$, $a_2 \in V_2$. Then the trace inner product is the unique inner product in $\mathcal{L}_{p_1,p_2}$ which satisfies the relation

$$< A, B >=< a_2 \otimes a_1, b_2 \otimes b_1 >= (a_2, b_2)_2(a_1, b_1)_1$$

for all $a_1, b_1 \in V_1$ and $a_2, b_2 \in V_2$.

As an example of multivariate linear regression model it will be considered a family of growth curve models with random effects (Lange and Laird [15], Reinsel [19, 20], Beganu [7, 8]) which consists in repeated measurements of m characteristics performed at p different moments of time on n individuals. Then the random matrix of responses is

$$Y = AB(X' \otimes I_m) + \Lambda(X' \otimes I_m) + E \tag{11}$$

where $A \in \mathcal{L}_{r,n}$ is the between-individuals design matrix and $X \in \mathcal{L}_{q,p}$ is the within-individual design matrix, which are known matrices of full column ranks $r < n$, $q \leq p$, respectively. $\Lambda \in \mathcal{L}_{qm,n}$ is a matrix of random effects and $E \in \mathcal{L}_{pm,n}$ is the random matrix of disturbances. The lines of Λ and E are random vectors independent of each other and between them and identically distributed with zero expected means and the same covariance matrices Σ_λ and Σ_e, respectively. It is assumed that Σ_λ and Σ_e are known symmetric positive definite matrices. The unknown parameters of this model are the regression coefficients $B \in \mathcal{L}_{qm,r}$.

The special form of the model (11) involves that the expected mean of the random matrix of observations is

$$E(Y) = AB(X' \otimes I_m) = \mu$$

and the covariance matrix is

$$\mathrm{cov}(\mathrm{vec}\, Y) = I_n \otimes V = \Sigma \tag{12}$$

where

$$V = (XX') \otimes \Sigma_\lambda + I_p \otimes \Sigma_e \tag{13}$$

lies in Θ, a set of $p_m \times p_n$ symmetric positive definite matrices such that $I_{pm} \in \Theta$.

Let $\Omega = \{AB(X' \times I_m) \mid B \in \mathcal{L}_{qm,r}\}$ be a linear manifold in $\mathcal{L}_{pm,n}$ and $\mathcal{X} \subset V_1 \subset R^{pm}$ and $\mathcal{A} \subset V_2 \subset R^n$ be the ranges of the design matrices $X \times I_m$ and A, respectively.

It is known that the OLSE of μ (see Lange and Laird [15], Reinsel [19, 20] given by

$$\hat{\mu}_{\mathrm{OLSE}} = A(A'A)^{-1}A'Y([X(X'X)^{-1}X'] \otimes I_m) \tag{14}$$

is identical to the BLUE of μ

$$\hat{\mu}_{\mathrm{BLUE}} = A(A'A)^{-1}A'YV^{-1}(X \otimes I_m)[(X' \otimes I_m)V^{-1}(X \otimes I_m)]^{-1}(X' \otimes I_m) \tag{15}$$

which exists in model (11) if and only if the linear manifold Ω is invariant under Σ (Eaton [11]), or the linear vector space $\mathcal{X}$ is invariant under V (Beganu [8]).

In order to prove that the equality between (14) and (15) holds using a coordinate-free approach, it is easy to notice that

$$\hat{\mu}_{\mathrm{OLSE}} = (P_A \copyright P_X)Y \tag{16}$$

is the unique element of Ω satisfying the relationship

$$< Y - (P_A \copyright P_X)Y, Z >=< Y - P_A Y P_X, Z >= 0$$

for all $Z \in \Omega$, where $P_A = A(A'A)^{-1}A'$ and $P_X = [X(X'X)^{-1}X'] \otimes I_m$ are the orthogonal projections onto $\mathcal{A}$ and $\mathcal{X}$ relative to the usual inner product on V_2 and V_1, respectively.

If the covariance operator corresponding to (12) is considered, the new inner product (Kruskal [14], Eaton [11]) denoted by

$$\ll U, Z \gg =< U, \Sigma^{-1} Z >=< U, (I_n \copyright V^{-1})Z >=< U, ZV^{-1} >$$

is defined on $\mathcal{L}_{pm,n}$ and the BLUE of μ

$$\hat{\mu}_{\mathrm{BLUE}} = (P_A \copyright Q_X)Y \tag{17}$$

is the unique element in Ω such that

$$\ll Y - (P_A ⓒ Q_X)Y, Z \gg = \ll Y - P_A Y Q_X^*, Z \gg = 0$$

for all $Z \in \Omega$. The linear operator on V_1 to $\mathcal{X}$

$$Q_X = (X \otimes I_m)[(X' \otimes I_m)V^{-1}(X \otimes I_m)]^{-1}(X' \otimes I_m)V^{-1}$$

is the orthogonal projection onto $\mathcal{X}$ with respect to $((u,z))_1 = (u, V^{-1}z)_1$ for all $u, z \in V_1$.

The estimators (14) and (15) expressed by the relations (16) and (17), respectively, will be used in the sequel to establish the equality (1) by means of Propositions 1, 2 and 3.

Replacing the orthogonal projections onto Ω relative to the two inner products defined on the real vector space $\mathcal{L}_{pm,n}$, we obtain that the first term in (5) is

$$\begin{aligned} \ll PU, PZ \gg &= < (P_A ⓒ P_X)U, (I_n ⓒ V^{-1}) \circ (P_A ⓒ P_X)Z > \\ &= < U, (P_A ⓒ P_X V^{-1} P_X)Z > \end{aligned}$$

and the second term is

$$\ll PU, Z \gg = < (P_A ⓒ P_X)U, (I_n ⓒ V^{-1})Z > = < U, (P_A ⓒ P_X V^{-1})Z >$$

for all $U, Z \in \mathcal{L}_{pm,n}$.

Since $U, Z \in \mathcal{L}_{pm,n}$ has ranks one, the former equality can be written as

$$< u_2 \otimes u_1, P_A z_2 \otimes z_1 P_X V^{-1} P_X > = < u_2 \otimes u_1, P_A z_2 \otimes z_1 V^{-1} P_X >$$

for all $u_1, z_1 \in V_1$ and $u_2, z_2 \in V_2$, which is equivalent to

$$(u_1, P_X V^{-1} P_X z_1)_1 = (u_1, P_X V^{-1} z_1)_1 \tag{18}$$

for all $u_1, z_1 \in V_1$.

In a similar way the conditions (6) and (9) equivalently expressed by the equalities (8) and (10), respectively, lead to the corresponding necessary and sufficient conditions for the equality (1) involving the othogonal projections P_X and Q_X onto $\mathcal{X}$ as follows:

$$(u_1, V^{-1} Q_X^* z_1)_1 = (u_1, P_X V^{-1} Q_X^* z_1)_1 \tag{19}$$

and

$$(u_1, P_X V^{-1} P_X z_1)_1 = (u_1, V^{-1} P_X z_1)_1 \tag{20}$$

for all $u_1, z_1 \in V_1$.

The results obtained in relations (18), (19) and (20) can be formulated in

Proposition 4 *The necessary and sufficient conditions for $\hat{\mu}_{\mathrm{OLSE}}$ to be $\hat{\mu}_{\mathrm{BLUE}}$ in model* (11) *are assigned to the real vector space $\mathcal{X}$ spanned by the columns of the within-individuals design matrix $X \otimes I_m$ of the model. The relations* (18), (19) *and* (20) *are alternative forms of the condition that $\mathcal{X}$ is invariant under V (or under V^{-1}).*

The necessary and sufficient conditions in Theorems 1, 2 and 3 are verified by the model (1), hence the BLUE (10) of μ exists and is the OLSE (9). These two questions are equivalent in the multivariate growth-curve model (1) because of its special covariance structure.

Corollaries 1 and 2 assert that the necessary and sufficient conditions for the existence of $\hat{\mu}_{\mathrm{BLUE}}$ and its equality with $\hat{\mu}_{\mathrm{OLSE}}$ in model (1) are independent on the between-individuals design matrix of the model. The special form of the orthogonal projection onto Ω allows these conclusions.

References

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

[2] Baksalary, J. K. and Kala, R., (1976). Criteria for estimability in multivariate linear models, *Math. Operationsforsch. Statist. Ser. Statist.*, **7**, 5–9.

[3] Baksalary, J. K. and van Eijnsbergen, A. C, (1988). A comparison of two criteria for ordinary least squares estimators to be best linear unbiased estimators, *Amer. Statist.*, **42**, 205–208.

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

[5] Beganu, G., (2003). The existence conditions of the best linear unbiased estimators of the fixed effects, *Econom. Comput. Econom. Cybernet. Studies and Research*, **36**, 95–102.

[6] Beganu, G., (2005). On Gram-Schmidt orthogonalizing process of design matrices in linear models as estimating procedure of covariance components, *Rev. R. Acad. Cien., Serie A Mat.*, **99**, 187–194.

[7] Beganu, G. 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.*, (to appear).

[8] Beganu, G. The best linear unbiased estimators of regression coefficients in a multivariate growth-curve model. A coordinate-free approach, *Rev. Colombiana de Matematicas*, (to appear).

[9] Drygas, H., (1972). A note on Gauss-Markov estimation in multivariate linear models, *Colloquia Matematica Societatis Janos Bolyai*, European Meeting of Statisticians, Budapest, **9**, 181–190.

[10] Drygas, H., (1975). Estimation and prediction for linear models in general spaces, *Math. Operationsforsch. Statist., ser. Statist.*, **8**, 301–324.

[11] Eaton, M. L., (1970). Gauss-Markov estimation for multivariate linear models: A coordinate-free approach, *Ann. Math. Statist.*, **2**, 528–538.

[12] Haberman, S. I., (1975). How much do Gauss-Markov and least squares estimates differ? A coordinate-free approach, *Ann. Statist.*, **3**, 982–990.

[13] Halmos, P. R., (1958). *Finite Dimensional Vector Spaces*, 2nd ed., Van Nostrand, Princeton, New Jersey.

[14] Kruskal, W., (1968). When are Gauss-Markov and least squares estimators identical? A coordinate-free approach, *Ann. Math. Statist.*, **39**, 70–75.

[15] 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.

[16] Puntanen, S. and Styan, G. P. H., (1989). The equality of the ordinary least squares estimator and the best linear unbiased estimator, *Amer. Statist.*, **43**, 153–161.

[17] Puntanen, S., Styan, G. P. H. and Tian, Y., (2005). Three rank formulas associated with the covariance matrices of the BLUE and the OLSE in the general linear model, *Econometric Theory*, **21**, 659–664.

[18] Rao, C. R., (1973). *Linear Statistical Inference and Its Applications* (2nd ed.), John Wiley, New York.

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

[20] Reinsel, C. G., (1984). Estimation and prediction in a multivariate random effects generalized linear models, *J. Amer. Statist., Assoc.*, **79**, 406–414.

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

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

[23] Zyskind, G., (1967). On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models, *Ann. Math. Statist.*, **38**, 1092–1109.

[24] Zyskind, G. and Martin, F. B., (1969). On best linear estimation and a general Gauss-Markoff theorem in linear models with arbitrary non-negative structure, *SIAM J. Appl. Math.*, **17**, 1190–1202.

Gabriela Beganu
Department of Mathematics
Academy of Economic Studies
Piaţa Romană, nr. 6
010374 Bucharest, Romania
gabriela_beganu@yahoo.com