Aplikace matematiky

Gejza Wimmer

On equivalence problem in linear regression models. I. BLUE of the mean value

*Aplikace matematiky*, Vol. 25 (1980), No. 6, 417–422

Persistent URL: http://dml.cz/dmlcz/103880

# ON EQUIVALENCE PROBLEM IN LINEAR REGRESSION MODELS

## Part I. BLUE OF THE MEAN VALUE

GEJZA WIMMER

(Received October 19, 1978)

## INTRODUCTION

If the stastic calculations of the linear estimation of regression parameters in regression models are extensive, the problem of checking and numerical stability of the obtained results arises. This article is a contribution to the solution of this problem.

Let us have the general regression model $(Y_{n,1}, X_{n,m}\beta_{m,1}, \sigma^2 V_{n,n})$, i.e. the random variable Y has the mean value $\mathbf{E}(Y) = X\beta$ and the covariance matrix $\sigma^2 V$, where $\beta$ is a vector of unknown parameters, X and V known matrices (not necessarily of full rank). Parameter $\sigma^2$ is a known or unknown scalar factor. If $p \in \mathcal{R}^m$ ($m$ dimensional vector space), then the linear function $p'\beta$ of parameters β is unbiased estimable iff $p \in \boldsymbol{\mu}(X')$, where $\boldsymbol{\mu}(X')$ is the vector space spanned by the columns of the matrix $X'$ (the transpose of X). The best linear unbiased estimation (BLUE) (i.e. unbiased estimation with minimal variance) of $p'\beta$ is $(p'\beta)^\wedge$ and we can obtain this BLUE in several different ways.

a) In "PANDORA-BOX" matrix

$$(1) \qquad \begin{pmatrix} V & X \\ X' & 0 \end{pmatrix}^- = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{pmatrix},$$

the BLUE satisfies

$$(2) \qquad (p'\beta)^\wedge = p'C_2'Y = p'C_3Y\,,$$

(see (4i. 3.2) in [2]), $A^-$ is a $g$-inverse of A, i.e. a matrix satisfying the relation $AA^-A = A$.

b) For an arbitrary choice of the minimum V-seminorm $g$-inverse $(X')^-_{m(V)}$ of the matrix $X'$, i.e. a matrix defined by the relations $X'(X')^{-}_{m(V)}X' = X'$, $((X')^{-}_{m(V)}X')'\,V = V(X)'^{-}_{m(V)}X'$, the BLUE is

$$(p'\beta)^{\wedge} = p'[(X')^{-}_{m(V)}]'\,Y \tag{3}$$

(see [3] p. 148).

c) Model $(Y, X\beta, \sigma^2 V)$ can be transformed into the model with conditions on the parameter and this model into the regular regression model. In this model there exists a unique BLUE of the unknown parameters. By the transformation of this estimation we can obtain the Blue $(p'\beta)^{\wedge}$ in the general model.

d) For the solution $\hat{\beta}$ of consistent normal equations (see [1] and [4])

$$X'MX\hat{\beta} = X'MY\,, \tag{4}$$

where

$$\begin{gathered} M = (V + XUX')^{-} + K\,, \\ \mathbf{R}(X'MX) = \mathbf{R}(X') \\ \boldsymbol{\mu}(V \vdots X) = \boldsymbol{\mu}(V + XUX') = \boldsymbol{\mu}(V + XU'X') \\ VK'X = 0 \quad X'KX = 0\,, \end{gathered} \tag{5}$$

it holds that the BLUE is

$$(p'\beta)^{\wedge} = p'\hat{\beta}\,. \tag{6}$$

$\mathbf{R}(X'MX)$ is the rank of the matrix $X'MX$.

Remark. Conditions (5) are necessary and sufficient for M, U, K in order that (6) might hold for every estimable function $p'\beta$.

We see that estimations $(p'\beta)^{\wedge}$ according to a), b), c) and d) are statistically equivalent (all of them are BLUE). A serious question arises if they are equivalent in realizations. It is an important question if we want to check the calculations in different ways on computers. This question has some importance also in the determination of numerical stability of statical estimation algorithms. Contradictions in results obtained in different ways lead to mistrust of the results and make the practical use of them impossible.

In the sequel we prove that all of above mentioned estimations coincide a.e. in realizations, i.e. that a numerical contradiction resulting from calculations in two different ways indicates a mistake in the program or numerical instability of the computation. These results may be also used when studying numerical stability of linear statistical estimation algorithms.

**Lemma 1.** *It holds*

$$\boldsymbol{P}\{\mathrm{Y} - \boldsymbol{E}(\mathrm{Y}) \in \boldsymbol{\mu}(\mathrm{V})\} = 1 .$$

Proof of this well known fact is e.g. in [2] p. 297. From Lemma 1 we have

$$\boldsymbol{P}\{\mathrm{Y} = \mathrm{Xa} + \mathrm{Vb},\ \mathrm{a} \in \mathscr{R}^m,\ \mathrm{b} \in \mathscr{R}^n\} = 1 . \tag{7}$$

**Lemma 2.** *If* $(\mathrm{X}')^-_{1,m(\mathrm{V})}$ *and* $(\mathrm{X}')^-_{2,m(\mathrm{V})}$ *are two minimum V-seminorm g-inverses of the matrix* X', *then*

$$\mathrm{X}[(\mathrm{X}')^-_{1,m(\mathrm{V})} - (\mathrm{X}')^-_{2,m(\mathrm{V})}]'\ \mathrm{V}[(\mathrm{X}')^-_{1,m(\mathrm{V})} - (\mathrm{X}')^-_{2,m(\mathrm{V})}]\ \mathrm{X}' = 0 .$$

Proof is easily obtained from the definition of the $g$-inverse $(\mathrm{X}')^-_{m(\mathrm{V})}$.

If we define the norm $\|\cdot\|_{\mathrm{V}^-}$ in the space $\boldsymbol{\mu}(\mathrm{V})$ by

$$\|\mathrm{Vu}\|^2_{\mathrm{V}^-} = \mathrm{u}'\mathrm{VV}^-\mathrm{Vu} ,$$

(V is a p.s.d. matrix), then an arbitrary element u fulfils

$$\mathrm{X}\{[(\mathrm{X}')^-_{1,m(\mathrm{V})}]' - [(\mathrm{X}')^-_{2,m(\mathrm{V})}]'\}\ \mathrm{Vu} = \mathrm{V}[(\mathrm{X}')^-_{1,m(\mathrm{V})} - (\mathrm{X}')^-_{2,m(\mathrm{V})}]\ \mathrm{X}'\mathrm{u} \in \boldsymbol{\mu}(\mathrm{V}) ,$$

according to Lemma 2 it is

$$\|\mathrm{X}\{[(\mathrm{X}')^-_{1,m(\mathrm{V})}]' - [(\mathrm{X}')^-_{2,m(\mathrm{V})}]'\}\ \mathrm{Vu}\|_{\mathrm{V}^-} = 0$$

for any two $g$-inverses $(\mathrm{X}')^-_{1,m(\mathrm{V})}$, $(\mathrm{X}')^-_{2,m(\mathrm{V})}$ and that is why

$$\mathrm{X}[(\mathrm{X}')^-_{1,m(\mathrm{V})}]'\ \mathrm{V} = \mathrm{X}[(\mathrm{X}')^-_{2,m(\mathrm{V})}]'\ \mathrm{V} . \tag{8}$$

**Theorem 1.** *If* $\mathrm{p} \in \boldsymbol{\mu}(\mathrm{X}')$, *then* BLUE $\mathrm{p}'[(\mathrm{X}')^-_{m(\mathrm{V})}]'\ \mathrm{Y}$ *is invariant for any choice of* $(\mathrm{X}')^-_{m(\mathrm{V})}$ *with probability* 1.

Proof. If $\mathrm{p} \in \boldsymbol{\mu}(\mathrm{X}')$, then $\mathrm{p} = \mathrm{X}'\mathrm{u}$ and according to (7) and (8),

$$\mathrm{p}'[(\mathrm{X}')^-_{m(\mathrm{V})}]'\ \mathrm{Y} = \mathrm{u}'\mathrm{X}[(\mathrm{X}')^-_{m(\mathrm{V})}]'\ (\mathrm{Xa} + \mathrm{Vb}) = \mathrm{u}'\mathrm{Xa} + \mathrm{u}'\mathrm{X}[(\mathrm{X}')^-_{m(\mathrm{V})}]'\ \mathrm{Vb}$$

with probability 1 and it does not depend on the choice of $(\mathrm{X}')^-_{m(\mathrm{V})}$. The proof is complete.

**Corollary 1.** *All the* BLUE *in* (3) *coincide numerically with probability* 1.

**Corollary 2.** *Because it is true that* $\mathrm{C}_2$ *is one choice of* $(\mathrm{X}')^-_{m(\mathrm{V})}$ *in* (1), $\mathrm{C}_3$ *is one choice* $[(\mathrm{X}')^-_{m(\mathrm{V})}]'$ (*see* (4i. 1.3) *in* [2]), *the estimations* (2) *and* (3) *coincide numerically with probability* 1.

**Lemma 3.**

$$(X'MX)^- X'M$$

*is one choice of* $[(X')^-_{m(V)}]'$ *for an arbitrary choice of* $(X'MX)^-$, *where* M *is defined in* d).

Proof. It is easy to see that $(X'MX)^- X'M$ is a $g$-inverse $[(X')^-]'$ for an arbitrary choice of $(X'MX)^-$.

The equation

$$X'MV(I - (X')^- X') = X'M(V + XUX' - XUX')(I - (X')^- X') =$$
$$= X'[(V + XUX')^- + K](V + XUX')(I - (X')^- X') -$$
$$- X'[(V + XUX')^- + K] XUX'(I - (X')^- X') = X'(I - (X')^- X') -$$
$$- X'[(V + XUX')^- + K] XUX'(I - (X')^- X') = 0$$

is true for an arbitrary $(X')^-$, therefore we obtain for $M'X[(X'MX)^-]'$

$$X'MV = X'MV(X')^- X' = X'MVM'X[(X'MX)^-]' X' . \tag{9}$$

We have the equation

$$X(X'MX)^- X'MV = X(X'MX)^- X'MVM'X[(X'MX)^-]' X' =$$
$$= VM'X[(X'MX)^-]' X' .$$

According to the definition of $(X')^-_{m(V)}$ the proof is complete.

**Corollary 3.** BLUE $(p'\beta)^\wedge = p'\hat{\beta}$ *obtained by solving* (4) *is numerically equivalent in realiziations with* (2) *and* (3).

How to compute the BLUE $(p'\beta)^\wedge$ for $p \in \mu(X')$ by transforming the general model to the regular one.

If we denote by N the matrix of order $n \times (n - \mathbf{R}(V))$ with rank $n - \mathbf{R}(V)$ such that $N'V = O$, $V = JJ'$ where $J_{n \times \mathbf{R}(V)}$ i.e. $\mathbf{R}(V) = \mathbf{R}(J)$ and $F' = (J'J)^{-1} J'$, then it holds

$$F'J = I_{\mathbf{R}(V),\mathbf{R}(V)}$$
$$F'N = 0$$

and $\begin{pmatrix} N' \\ F' \end{pmatrix}$ is a nonsingular matrix. Let us take an arbitrary but fixed $g$-inverse $(N'X)^-$.

The random vector $F'(I - X(N'X)^- N')\, Y$ has the mean value $F'X(I - (N'X)^- . N'X)\,\eta$ $(\eta \in \mathscr{R}^m)$ and the covariance matrix $\sigma^2 I$. The matrix $F'X(I - (N'X)^- N'X)$ can be written (see (1c. 3.11) in [2]) as

$$F'X(I - (N'X)^- N'X) = P\Lambda Q' , \tag{10}$$

where P, $\Lambda$ and Q are matrices of full rank and $P'P = Q'Q = I$. $\Lambda$ is a diagonal matrix of order

$\boldsymbol{R}(F'X(I - (N'X)^- N'X)) \times \boldsymbol{R}(F'X(I - (N'X)^- N'X))$, P of order

$\boldsymbol{R}(V) \times \boldsymbol{R}(F'X(I - (N'X)^- N'X))$ and Q of order

$m \times \boldsymbol{R}(F'X(I - (N'X)^- N'X))$.

We see that the random vector $F'(I - X(N'X)^- N')Y$ has the mean value $P\gamma$, where $\gamma = \Lambda Q'\eta \in \mathscr{R}^{\boldsymbol{R}(F'X(I-(N'X)^-N'X))}$ and the covariance matrix $\sigma^2 I$. We have the regular regression model $(F'(I - X(N'X)^- N')Y, P\gamma, \sigma^2 I)$, i.e. the dimension of the random vector $F'(I - X(N'X)^- N')Y$ is $\dim F'(I - X(N'X)^- N')Y = \boldsymbol{R}(V) \geqq \geqq \dim \gamma = \boldsymbol{R}(F'X(I - (N'X)^- N'X))$ and P is a matrix of full rank. In this regular model there exists a unique BLUE for every linear functional $q'\gamma$ and it is $(q'\gamma)\hat{} = q'\hat{\gamma}$ where (according to [3] p. 140) $\hat{\gamma} = P'F'(I - X(N'X)^- N')Y$.

**Lemma 4.**

$$(p'\beta)\hat{} = p'(N'X)^- N'Y + p'(I - (N'X)^- N'X) Q\Lambda^{-1}\hat{\gamma}$$

*is the* BLUE *for* $p'\beta$ *if* $p \in \boldsymbol{\mu}(X')$.

Proof. We shall show that the matrix

$$(N'X)^- N' + (I - (N'X)^- N'X) Q\Lambda^{-1}P'F'(I - X(N'X)^- N')$$

is one choice of the $g$-inverse $[(X')^-_{m(V)}]'$.

According to (10) and to the equations $\boldsymbol{\mu}((I - (N'X)^- N'X)' X') = \boldsymbol{\mu}((I - - (N'X)^- N'X)' X'(F \vdots N)) = \boldsymbol{\mu}((I - (N'X)^- N'X)' X'F \vdots (I - (N'X)^- N'X)' X'N) = = \boldsymbol{\mu}((I - (N'X)^- N'X)' X'F) = \boldsymbol{\mu}(Q\Lambda P')$ (i.e. there exists a matrix K such that $X(I - (N'X)^- N'X) = KP\Lambda Q'$) it holds

$$X(I - (N'X)^- N'X) Q\Lambda^{-1}P'F'X(I - (N'X)^- N'X) =$$
$$= KP\Lambda Q'Q\Lambda^{-1}P'P\Lambda Q' = X(I - (N'X)^- N'X)$$

and we have the equation

$$X\{(N'X)^- N' + (I - (N'X)^- N'X) Q\Lambda^{-1}P'F'(I - X(N'X)^- N')\} X =$$
$$= X(N'X)^- N'X + X(I - (N'X)^- N'X) = X.$$

Further, it holds

(11) $$X\{(N'X)^- N' + (I - (N'X)^- N'X) Q\Lambda^{-1}P'F'(I - (N'X)^- N')\} V =$$
$$= X(I - (N'X)^- N'X) Q\Lambda^{-1}P'F'V.$$

We shall show that the matrix (11) is symmetric. From the equation

$$\begin{pmatrix} P\Lambda^{-1}Q'Q\Lambda P' & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} P\Lambda Q'Q\Lambda^{-1}P' & 0 \\ 0 & 0 \end{pmatrix}$$

we have

$$\begin{pmatrix} P\Lambda^{-1}Q'(I-(N'X)^{-}N'X)'X' \\ 0 \end{pmatrix}(F \vdots N) =$$

$$= \begin{pmatrix} F'X(I-(N'X)^{-}N'X)Q\Lambda^{-\prime}P' \\ 0 \end{pmatrix}(I \vdots 0)$$

$$\begin{pmatrix} F' \\ N' \end{pmatrix} VFP\Lambda^{-1}Q'(I-(N'X)^{-}N'X)'X'(F \vdots N) =$$

$$= \begin{pmatrix} F' \\ N' \end{pmatrix} X(I-(N'X)^{-}N'X)Q\Lambda^{-1}P'F'V(F \vdots N)$$

and by virtue of the definition of $(X')^{-}_{m(V)}$ and (3) the lemma is proved.

**Corollary 4.** *For the* BLUE $(p'\beta)^{\wedge}$ *obtained by the reduction of the general model* $(Y, X\beta, \sigma^2 V)$ *to the regular case* $(F'(I - X(N'X)^{-}N')Y, P\gamma, \sigma^2 I)$ *it holds*

$$(p'\beta)^{\wedge} = p'(N'X)^{-}N'Y + p'(I-(N'X)^{-}N'X)Q\Lambda^{-1}\hat{\gamma} = p'[(X')^{-}_{m(V)}]'Y$$

*and therefore this* BLUE *is numerically equivalent with the estimation* (2), (3) *and* (6) *with probability* 1.

## CONCLUSIONS

The aim of this part was to show the numerical equivalences of the basic statistical estimation algorithms. These equivalences may increase the reliability of numerical calculations on computers. Corollaries 1, 2, 3 and 4 of this article are important for the practical use.

### *References*


[1] *C. R. Rao:* Unified Theory of Linear Estimation. Sankhyā, Vol. 33 1971, pp. 371—394.
[2] *C. R. Rao:* Linear Statistical Inference and Its Applications. John Wiley, N. York 1973.
[3] *C. R. Rao, S. K. Mitra:* Generalized Inverse of Matrices and Its Applications. John Wiley, N. York 1971.
[4] *C. R. Rao:* Corrigenda. Sankhyā A, Vol. 34 1972 p. 477.