Sankhyā: The Indian Journal of Statistics
1997, Volume 59, Series A, Pt. 2, pp. 175-178

WHEN DO LINEAR TRANSFORMS OF ORDINARY LEAST SQUARES AND GAUSS-MARKOV ESTIMATOR COINCIDE ?

By JÜRGEN GROß
and
GÖTZ TRENKLER*
University of Dortmund

SUMMARY. This paper investigates equality of ordinary least squares and Gauss-Markov estimator when we are interested in a linear transform of the parameter vector. Some applications to the estimation of subvectors are given.

1. Introduction

For a given matrix $\mathbf{A}$ its range is denoted by $\mathcal{R}(\mathbf{A})$, and its nullspace by $\mathcal{N}(\mathbf{A})$. By $\mathbf{A}^{\perp}$ we denote any matrix of maximal rank such that $\mathbf{A}^T\mathbf{A}^{\perp} = \mathbf{0}$, where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$.

Consider the linear regression model

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}, \qquad \ldots(1)$$

where $E(\mathbf{u}) = \mathbf{0}$, $Cov(\mathbf{u}) = \mathbf{V}$. We assume that $\mathbf{X} \in I\!R^{n\times p}$ is of full column rank, and $\mathbf{V} \in I\!R^{n\times n}$ is positive definite. To estimate the parameter vector β we may use the two competing statistics

$$\mathbf{b}(\mathbf{y}) = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \qquad \ldots(2)$$

and

$$\widehat{\beta}(\mathbf{y}) = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}. \qquad \ldots(3)$$

Suppose now we wish to estimate the linear transform $\mathbf{B}\beta$, where $\mathbf{B}$ is any

Paper received. October 1996.
AMS (1990) *subject classification.* 62J05.
Key words and phrases. Linear regression model, ordinary least squares estimator, Gauss-Markov estimator.
* Support by the Deutsche Forschungsgemeinschaft under grant Tr 253/2-1/2-2 is gratefully acknowledged.

given $q \times p$ matrix. Then $\mathbf{Bb(y)}$ and $\mathbf{B}\widehat{\beta}(\mathbf{y})$ are unbiased and $\mathbf{B}\widehat{\beta}(\mathbf{y})$ is optimal within the class of linear statistics with respect to the covariance matrix criterion (Gauss-Markov-Theorem). Nevertheless there might exist situations where both statistics coincide. Krämer (1980) has identified all $\mathbf{y}$ values such that $\mathbf{b(y)}$ and $\widehat{\beta}(\mathbf{y})$ are equal. Following his approach subsequently we investigate the more general problem of identification of all vectors $\mathbf{y}$ such that $\mathbf{Bb(y)} = \mathbf{B}\widehat{\beta}(\mathbf{y})$.

For this purpose define the following matrices: $\mathbf{D} = \mathbf{BX}^+$, $\mathbf{P} = \mathbf{XX}^+$ and $\mathbf{R} = \mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}$, where $\mathbf{X}^+ = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ is the Moore–Penrose inverse of $\mathbf{X}$. Obviously we have $\mathbf{Bb(y)} = \mathbf{B}\widehat{\beta}(\mathbf{y})$ if and only if $\mathbf{y} \in \mathcal{N}[\mathbf{D}(\mathbf{P}-\mathbf{R})]$.

2. Results

To analyze $\mathcal{E} = \mathcal{N}[\mathbf{D}(\mathbf{P}-\mathbf{R})]$, which is the set of all $\mathbf{y}$–vectors where $\mathbf{Bb(y)} = \mathbf{B}\widehat{\beta}(\mathbf{y})$, we need the subsequent lemma.

LEMMA 1. *The following identity holds*

$$\mathcal{N}(\mathbf{DR}) = \mathcal{R}(\mathbf{VX}^\perp) \oplus [\mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{D})].$$

PROOF. Before showing the asserted equality note that according to Rao (1974) we have $\mathcal{R}(\mathbf{X} : \mathbf{V}) = \mathcal{R}(\mathbf{X} : \mathbf{VX}^\perp)$ and $\mathcal{R}(\mathbf{X}) \cap \mathcal{R}(\mathbf{VX}^\perp) = \{\mathbf{0}\}$. Furthermore take into account that $\mathcal{R}(\mathbf{X} : \mathbf{V}) = I\!R^n$.

First we demonstrate the inclusion "$\subseteq$". Every vector $\mathbf{y} \in I\!R^n = \mathcal{R}(\mathbf{X} : \mathbf{V}) = \mathcal{R}(\mathbf{X} : \mathbf{VX}^\perp)$ may be written as $\mathbf{y} = \mathbf{VX}^\perp\mathbf{a} + \mathbf{b}$, where $\mathbf{a}$, $\mathbf{b}$ are vectors, and $\mathbf{b} \in \mathcal{R}(\mathbf{X})$. If in addition $\mathbf{y} \in \mathcal{N}(\mathbf{DR})$ we get $\mathbf{0} = \mathbf{DRy} = \mathbf{DRVX}^\perp\mathbf{a} + \mathbf{DRb} = \mathbf{Db}$. Hence $\mathbf{y} \in \mathcal{R}(\mathbf{VX}^\perp) \oplus [\mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{D})]$.

To show the reverse inclusion let $\mathbf{y} = \mathbf{VX}^\perp\mathbf{a} + \mathbf{b}$ for some vectors $\mathbf{a}$ and $\mathbf{b}$ such that $\mathbf{b} \in \mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{D})$. Then we conclude $\mathbf{RVX}^\perp = \mathbf{0}$, $\mathbf{Rb} = \mathbf{b}$ and $\mathbf{Db} = \mathbf{0}$ which yields $\mathbf{DRy} = \mathbf{0}$. □

Let us now proceed to characterize the class $\mathcal{E} = \mathcal{N}[\mathbf{D}(\mathbf{P}-\mathbf{R})]$ consisting of all $\mathbf{y}$–vectors where $\mathbf{Bb(y)} = \mathbf{B}\widehat{\beta}(\mathbf{y})$.

THEOREM 1. *The following identity holds*

$$\mathcal{E} = \mathcal{R}(\mathbf{X}) \oplus [\mathcal{R}^\perp(\mathbf{X}) \cap \mathcal{N}(\mathbf{DR})].$$

PROOF. "$\subseteq$" Let $\mathbf{y} \in \mathcal{E}$. Then $\mathbf{y} \in I\!R^n = \mathcal{R}(\mathbf{X}) \oplus \mathcal{R}^\perp(\mathbf{X})$, i.e. $\mathbf{y} = \mathbf{a} + \mathbf{b}$, $\mathbf{a} \in \mathcal{R}(\mathbf{X})$, $\mathbf{b} \in \mathcal{R}^\perp(\mathbf{X})$. Then we get $\mathbf{DPy} = \mathbf{DPa}$ and $\mathbf{DRy} = \mathbf{DRa} + \mathbf{DRb} = \mathbf{Da} + \mathbf{DRb}$. Since $\mathbf{y} \in \mathcal{E}$ we have $\mathbf{D}(\mathbf{P}-\mathbf{R})\mathbf{y} = \mathbf{0}$ which implies $\mathbf{DRb} = \mathbf{0}$, i.e. $\mathbf{y} \in \mathcal{R}(\mathbf{X}) \oplus [\mathcal{R}^\perp(\mathbf{X}) \cap \mathcal{N}(\mathbf{DR})]$.

"$\supseteq$" Let $\mathbf{y} = \mathbf{Xa} + \mathbf{b}$, where $\mathbf{a}$, $\mathbf{b}$ are vectors such that $\mathbf{b} \in \mathcal{R}^\perp(\mathbf{X}) \cap \mathcal{N}(\mathbf{DR})$. Then $\mathbf{X}^T\mathbf{b} = \mathbf{0}$ and $\mathbf{DRb} = \mathbf{0}$. Since $(\mathbf{P}-\mathbf{R})\mathbf{X} = \mathbf{0}$ we obtain $\mathbf{D}(\mathbf{P}-\mathbf{R})\mathbf{y} = \mathbf{D}(\mathbf{P}-\mathbf{R})\mathbf{Xa} + \mathbf{D}(\mathbf{P}-\mathbf{R})\mathbf{b} = \mathbf{D}(\mathbf{P}-\mathbf{R})\mathbf{b} = \mathbf{DPb} - \mathbf{DRb} = \mathbf{0} - \mathbf{0} = \mathbf{0}$. □

COROLLARY 1. *Let* $\mathbf{B} = \mathbf{I}$. *Then we get*

$$\mathcal{E} = \mathcal{R}(\mathbf{X}) \oplus [\mathcal{R}^{\perp}(\mathbf{X}) \cap \mathcal{R}(\mathbf{V}\mathbf{X}^{\perp})],$$

which is Krämer's (1980) result.

Let us now turn our attention to the case when $\mathbf{X}$ is partitioned as $\mathbf{X} = (\mathbf{X}_1 : \mathbf{X}_2)$ and $\mathbf{B} = (\mathbf{0} : \mathbf{I})$, i.e. we wish to estimate the lower part of $\boldsymbol{\beta}$. Then $\mathbf{X}^+$ may be partitioned accordingly as $\mathbf{X}^+ = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix}$, and Theorem 1 yields

$$\mathcal{E} = \mathcal{R}(\mathbf{X}) \oplus \mathcal{R}^{\perp}(\mathbf{X}) \cap [\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{Z}_2)].$$

To simplify this expression we need the following result:

LEMMA 2.

$$\mathcal{R}(\mathbf{X}) \cap \mathcal{N}(\mathbf{Z}_2) = \mathcal{R}(\mathbf{X}_1).$$

PROOF. Consider the identity

$$\mathbf{X}^+\mathbf{X} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix} (\mathbf{X}_1 : \mathbf{X}_2) = \begin{pmatrix} \mathbf{Z}_1\mathbf{X}_1 & \mathbf{Z}_1\mathbf{X}_2 \\ \mathbf{Z}_2\mathbf{X}_1 & \mathbf{Z}_2\mathbf{X}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

Hence $\mathbf{Z}_2\mathbf{X}_1 = \mathbf{0}$, and consequently $\mathcal{R}(\mathbf{X}_1) \subseteq \mathcal{N}(\mathbf{Z}_2)$, i.e. $\mathcal{R}(\mathbf{X}_1) \subseteq \mathcal{N}(\mathbf{Z}_2) \cap \mathcal{R}(\mathbf{X})$. Conversely, if $\mathbf{y} \in \mathcal{N}(\mathbf{Z}_2) \cap \mathcal{R}(\mathbf{X})$ we have $\mathbf{Z}_2\mathbf{y} = \mathbf{0}$ and $\mathbf{y} = \mathbf{X}_1\mathbf{a} + \mathbf{X}_2\mathbf{b}$ for some vectors $\mathbf{a}$ and $\mathbf{b}$. From the above identity related to $\mathbf{X}^+\mathbf{X}$ we obtain $\mathbf{Z}_2\mathbf{y} = \mathbf{Z}_2\mathbf{X}_1\mathbf{a} + \mathbf{Z}_2\mathbf{X}_2\mathbf{b} = \mathbf{b} = \mathbf{0}$ such that $\mathbf{y} = \mathbf{X}_1\mathbf{a} \in \mathcal{R}(\mathbf{X}_1)$. □

From this our next theorem may be stated.

THEOREM 2. *If* $\mathbf{B} = (\mathbf{0} : \mathbf{I})$ *we have*

$$\mathcal{E} = \mathcal{R}(\mathbf{X}) \oplus \mathcal{R}^{\perp}(\mathbf{X}) \cap [\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)].$$

Since $\mathbb{R}^n = \mathcal{R}(\mathbf{X}) \oplus \mathcal{R}^{\perp}(\mathbf{X})$, we also get

COROLLARY 2. *Let* $\mathbf{B} = (\mathbf{0} : \mathbf{I})$. *Then* $\mathcal{E} = \mathbb{R}^n$ *if and only if*

$$\mathcal{R}^{\perp}(\mathbf{X}) \subseteq [\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)].$$

Corollary 2 can be modified by observing that $\mathcal{R}^{\perp}(\mathbf{X}) = \mathcal{R}(\mathbf{X}^{\perp})$, and that the assertion of Corollary 2 is equivalent to

$$\mathcal{R}(\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)] \subseteq [\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_2)].$$

Since however $n - p = rk(\mathbf{X}^{\perp}) = rk(\mathbf{V}\mathbf{X}^{\perp})$ we conclude $rk(\mathbf{X}_1 : \mathbf{X}^{\perp}) = rk(\mathbf{X}_1 : \mathbf{V}\mathbf{X}^{\perp})$, which also gives

$$[\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)] \subseteq [\mathcal{R}(\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)]$$

or, equivalently

$$\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \subseteq [\mathcal{R}(\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)].$$

This gives the result achieved by Krämer *et al.* (1996):

COROLLARY 3. *Let* $\mathbf{B} = (\mathbf{0} : \mathbf{I})$. *Then* $\mathcal{E} = \mathbb{R}^n$ *if and only if*

$$\mathcal{R}(\mathbf{V}\mathbf{X}^{\perp}) \subseteq [\mathcal{R}(\mathbf{X}^{\perp}) \oplus \mathcal{R}(\mathbf{X}_1)].$$

References

KRÄMER, W. (1980). A note on the equality of ordinary least squares and Gauss–Markov estimates in the general linear model. *Sankhyā*, Ser. A, **42**, 130 – 131.

KRÄMER, W.; R. BARTELS and D.G. FIEBIG (1996). Another twist on the equality of OLS and GLS. *Statistical Papers* **37**, 277 – 281.

RAO, C.R. (1974). Projectors, generalized inverses and the BLUE's. *J. Roy. Statist. Soc.*, Ser. B, **36**, 442 – 448.

DEPARTMENT OF STATISTICS
UNIVERSITY OF DORTMUND
VOGELPOTHSWEG 87
D-44221 DORTMUND
GERMANY
gross@amadeus.statistik.uni-dortmund.de