

Equality of two blues and ridge-type estimates

Friedrich Pukelsheim

Angaben zur Veröffentlichung / Publication details:

Pukelsheim, Friedrich. 1977. "Equality of two blues and ridge-type estimates."
Communications in Statistics - Theory and Methods 6 (7): 603–10.
<https://doi.org/10.1080/03610927708827518>.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

Deutsches Urheberrecht

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



EQUALITY OF TWO BLUES AND RIDGE-TYPE ESTIMATES

Friedrich Pukelsheim

Institut für Mathematische Stochastik der Universität
D-7800 Freiburg im Breisgau, Fed. Rep. of Germany

ABSTRACT

Equality is shown of the g -inverse and Moore-Penrose inverse representation of the BLUE in the general linear model. The proof is based on a matrix identity which allows also to establish a functional relationship between the BLUE and Ridge-type estimates.

1. INTRODUCTION

The present communication focuses on some computational properties of the matrices that appear in BLUE and Ridge-type estimation in linear model theory. In Section 3 we shortly define what now we loosely call Ridge-type estimates, for its statistical import, however, the reader is referred to Hoerl & Kennard (1970), Rao (1973, p.306), or Rolph (1976), the latter including many additional refer-

ences. Procedures for mean estimation are also useful for the estimation of variance components, see Pukelsheim (1976).

Consider the general linear model

$$\mathbb{E}Y = Xb, \quad \mathbb{D}Y = \sigma^2 V^2, \quad (1)$$

where Y is an \mathbb{R}^n -valued random vector, X is a known real $n \times p$ matrix, and V^2 is a known dispersion matrix written as the square of its unique nonnegative definite symmetric square root V . Interest concentrates on linear estimators $\hat{b}Y$ for the vector parameter b , and on appropriate justifications which $p \times n$ matrix \hat{b} that is determining the estimator is to be chosen.

Section 2 deals with the g -inverse and the Moore-Penrose inverse representation of the BLUE. The class of all those matrices \hat{b} leading to BLUEs $q'\hat{b}Y$ for all estimable linear forms $q'b$, $q \in \mathbb{R}^p$, has been given two different representations by Albert (1973, p.184):

$$X^+(I - V(MV)^+) + Z(M - MV(MV)^+), \quad M = I - XX^+, \quad (2)$$

and by Mitra & Moore (1973, p.141):

$$(X'(V^2 + XX')^{-1}X)^{-1}X'(V^2 + XX')^{-1}. \quad (3)$$

The multiplicity is generated in (2) by the arbitrariness of the $p \times n$ matrix Z , and in (3) by the choice of the g -inverses. Mitra & Moore (1973, p.142) proved that

$$B := X^+(I - V(MV)^+) \quad (4)$$

is in the class (3); Proposition 1 below states more exactly that B is equal to the Moore-Penrose version in (3). Thus the naturally distinguished matrices in (2) and (3) coincide.

Section 3 turns to Ridge-type estimates since the term $X'(V^2 + XX')^+$ not only arises in BLUE theory as in (3) but is even more important for Ridge-type estimation, see Hoerl & Kennard (1970, p.57), Rao (1973, p.306), Rolph (1976, p.794). Proposition 2 shows how to compute the BLUE from Ridge-type estimates and vice versa; as a corollary we obtain various representations for Ridge-type estimates whose derivations follow easily from BLUE theory.

All proofs are collected in Section 4.

2. EQUALITY OF TWO BLUES

Proposition 1 proposes an answer to Albert's (1973, p.183) "question concerning the relationship between the matrices in (2) and (3)": Put $Z = 0$ in (2) and choose Moore-Penrose inverses in (3), and the resulting matrices are equal.

Proposition 1: $B = (X'(V^2 + XX')^+X)^+X'(V^2 + XX')^+.$

The proof is given in Section 4; its crucial step is the following matrix identity which follows from Cline's (1965, p.100) inverse for the sum of nonnegative definite matrices.

Lemma: $X'(V^2 + XX')^+ = (I + BV^2B')^{-1}B.$

Since the two terms in (2) are orthogonal with respect to the trace inner product of matrices, B is the shortest matrix in (2) and Proposition 1 has the

Corollary 1: $(X'(V^2 + XX')^+X)^+X'(V^2 + XX')^+$ is of minimum norm in the class (3) with respect to the Euclidean matrix norm.

In model (1) the variance component σ^2 is unknown; since, however, in $X^+(I - \sigma V(\sigma MV)^+)$ the σ cancels out, Proposition 1 gives rise to the further

Corollary 2: $B = (X'(\sigma^2 V^2 + XX')^+ X)^+(\sigma^2 V^2 + XX')^+$ for all $\sigma^2 > 0$.

It is obvious from (2) that the BLUE admits a unique linear representation if and only if $Z(M - MV(MV)^+) = 0$ for all Z . But $M - MV(MV)^+$ orthogonally projects onto the intersection of the nullspaces of X' and V^2 , which is the orthogonal complement of $\text{range } X + \text{range } V^2$, where range means column space. Thus we finally get the

Corollary 3: $B = (X'(V^2 + XX')^{-} X')^{-} X'(V^2 + XX')^{-}$ for all choices of g-inverses if and only if $V^2 + XX'$ is nonsingular. In this case $(V^2 + XX')^{-} = (V^2 + XX')^{-1}$.

Corollary 3 rather states that in (3) the versions of the g-inverses are not, in general, negligible in order to have equality with B .

While the estimator $q'b\hat{Y}$ for $q'b$, with \hat{b} from (2), need not be unbiased for all $q \in \mathbb{R}^p$, it is always the minimum variance - minimum bias - linear estimator (MV-MB-LE) for $q'b$, see Rao (1973, p.307). Particularly when unbiasedness is not possible, one is interested in alternative estimation procedures.

3. RIDGE-TYPE ESTIMATES

In model (1) the mean square error of a linear estimator $q'b\hat{Y}$ for $q'b$ is $\sigma^2 \|Vq\|^2 + \|(X'\hat{q} - q)'b\|^2$ with maximum value $\sigma^2 \|Vq\|^2 + \beta^2 \|X'\hat{q} - q\|^2$ when the vector parameter b

varies subject to $\|b\| \leq \rho$. Minimizing the maximal mean square error on the ball $\|b\| \leq \rho$ thus leads to the problem of minimizing

$$k\|\hat{V}\hat{q}\|^2 + \|X'\hat{q} - q\|^2, \quad k > 0. \quad (5)$$

The resulting estimators are $q'b_k^*Y$, where the defining equality for the $p \times n$ matrix b_k^* is, see Rao (1973, p.306),

$$b_k^* \cdot (kV^2 + XX') = X'. \quad (6)$$

In the present communication we call, per definition, b_k^*Y Ridge-type estimate for b whenever b_k^* solves (6).

The general solution to (6) is $b_k^* = X'(kV^2 + XX')^{-}$, and it follows from (3) that then $(b_k^*X)^{-}b_k^*Y$ is the MV-MB-LE for b , irrespective of the value of k . In particular, if $b_k^* = X'(kV^2 + XX')^{+}$, then $(b_k^*X)^{+}b_k^* = B$, by Proposition 1. Thus the MV-MB-LE may be computed when a Ridge-type estimate is given; Proposition 2 solves the converse problem.

Proposition 2: If $\hat{b}Y$ is a MV-MB-LE for b , i.e., \hat{b} is representable as in (2), and if $k > 0$, then $(I + k\hat{b}V^2\hat{b}')^{-1}\hat{b}Y$ is a Ridge-type estimate.

The proof follows from the Lemma and is given in Section 4. The functional relationship of \hat{b} and b_k^* may be used to derive alternative representations for b_k^* . The Aitken estimator $(X'V^{2+}X)^{+}X'V^{2+}Y$ is a MV-MB-LE if and only if $\text{range } X \subset \text{range } V^2$, see Zyskind (1975, p.658). The reader will then easily verify the

Corollary 4: $(kI + X'V^{2+}X)^{-1}X'V^{2+}Y$ is a Ridge-type estimate if and only if $\text{range } X \subset \text{range } V^2$.

The simple least squares estimator X^+Y is a MV-MB-LE if and only if $\text{range } V^2X \subset \text{range } X$, see Zyskind (1975, p.684), hence

Corollary 5: $(I + kX^+V^2X^+)^{-1}X^+Y$ is a Ridge-type estimate if and only if $\text{range } V^2X \subset \text{range } X$.

If $V^2 = I$ then Corollaries 4 and 5 apply and yield the representations (2.1) and (2.3) in Hoerl & Kennard (1970, p.57). We are now left with proving the Lemma and Propositions 1 and 2.

4. PROOFS

First, we prove the Lemma. Inverting the sum $V^2 + XX'$ with Cline's formula (1965, p.100) and some computation yield

$$X'(V^2 + XX')^+ = (I - BVKVX^+) \cdot B, \quad K = (I + VB'BV)^{-1}. \quad (7)$$

Now, $BVK = (I + BV^2B')^{-1}BV$, and $BV^2(VM)^+ = X^+V(I - (MV)^+MV) \cdot VM(VM)^+ = 0$. Hence

$$\begin{aligned} I - BVK \cdot VX^+ &= I - (I + BV^2B')^{-1}BV \cdot V \cdot (I - (VM)^+V + (VM)^+V) \cdot X^+ \\ &= I - (I + BV^2B')^{-1}(BV^2B' + 0 + I - I) \\ &= (I + BV^2B')^{-1}. \end{aligned} \quad (8)$$

The Lemma is then established by inserting (8) into (7).

Next, we prove Proposition 1. Clearly, $BX = X^+X$, and $B = X^+XB$. Using the Lemma, we obtain

$$\begin{aligned} (X'(V^2 + XX')^+X)^+X'(V^2 + XX')^+ \\ &= ((I + BV^2B')^{-1}BX)^+(I + BV^2B')^{-1}B \\ &= ((I + BV^2B')^{-1}X^+X)^+((I + BV^2B')^{-1}X^+X) \cdot B. \end{aligned}$$

Since the ranges of $X^+X(I + BV^2B')^{-1}$ and X^+X coincide, so do their projectors. Thus the last equalities may be continued $= X^+XB = B$, establishing Proposition 1.

Finally, we prove Proposition 2. The Lemma implies

$$\begin{aligned}(I + k \hat{b}V^2\hat{b}')^{-1}\hat{b} &= (I + k BV^2B')^{-1}B + Z(M - MV(MV)^+), \\ &= X'(kV^2 + XX')^+ + Z(M - MV(MV)^+),\end{aligned}$$

and postmultiplication with $kV^2 + XX'$ yields X' , and Proposition 2 is established.

ACKNOWLEDGEMENT

The author would like to thank Professor H. Witting for his encouragement and continued support while preparing this note.

BIBLIOGRAPHY

- Albert, A. (1973). The Gauss-Markov theorem for regression models with possibly singular covariances. SIAM J. Appl. Math. 24, 182 - 87.
- Cline, R.E. (1965). Representations for the generalized inverse of sums of matrices. SIAM J. Numer. Anal. 2, 99 - 114.
- Hoerl, A.E. and Kennard, R.W. (1970). Ridge regression: biased estimation for nonorthogonal problems. Technometrics 12, 55 - 67.
- Mitra, S.K. and Moore, B.J. (1973). Gauss-Markov estimation with an incorrect dispersion matrix. Sankhyā A 35, 139 - 52.
- Pukelsheim, F. (1976). Estimating variance components in linear models. J. Multivariate Anal. 6, 626 - 29.

- Rao, C.R. (1973). Linear Statistical Inference and Its Applications (2nd ed.). New York: John Wiley and Sons, Inc.
- Rolph, J.E. (1976). Choosing shrinkage estimators for regression problems. Commun. Statist. A 5, 789 - 802.
- Zyskind, G. (1975). Error structures, projections and conditional inverses in linear model theory. In A Survey of Statistical Design and Linear Models (J.N. Srivastava, Ed.). Amsterdam: North-Holland Publ. Co., pp. 647 - 63.