24

PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 5, Fasc. 2 (1985), p. 177-186

ON THE UNIFIED THEORY OF LEAST SQUARES [1]

BY

HILMAR DRYGAS (KASSEL)

Dedicated to the memory of my father,
Gerhard Drygas (1902-1983)

Abstract. This paper deals with the linear model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$ where deficiency of ranks is allowed for both X and V. It is investigated when artificially enlarging V to a matrix W may result in correct estimation and prediction formulae. This leads to condition on W and finally ends up in the Rao-Mitra approach of Unified Least Squares.

1. Introduction. We consider the simple linear model

$$y = X\beta + \sigma\varepsilon, \quad \mathrm{E}\,\varepsilon = 0, \quad \mathrm{E}\,\varepsilon\varepsilon' = \mathrm{Cov}\,\varepsilon = V,$$

implying $\mathrm{Cov}\, y = \sigma^2 \mathrm{Cov}\,\varepsilon = \sigma^2 V$. Here y is an observed random $n \times \times 1$-vector, X is a fixed $n \times k$-matrix, β is a $k \times 1$-parameter-vector, V a n.n.d. (non-negative definite) symmetric matrix and $\sigma > 0$ an unknown dispersion. The problem of estimating β has an easy solution if X has full column-rank k and V is regular. In this case

$$X\hat{\beta} = X(X'V^{-1}X)^{-1}X'V^{-1}y \tag{1.1}$$

is the Best Linear Unbiased Estimator (BLUE) of $X\beta$. This formula extends to the case where $\mathrm{im}(X) \subseteq \mathrm{im}(V)$ and X has less than full rank, if the occurring inverses are replaced by generalized inverses (g-inverses).

If, however, the condition $\mathrm{im}(X) \subseteq \mathrm{im}(V)$ is not met, a BLUE can no longer be obtained via the (modified) Aitken formula (1.1). To overcome the problem of deficient rank, Rao and Mitra have designed a method called the

[1] A previous version of this paper has been presented at the European Meeting of Statisticians, Wrocław, Poland, 31. 8.-4. 9. 81.

"Unified theory of least squares". This method has been elaborated in a series of papers (Mitra [7], Mitra and Moore [8], Rao and Mitra [10], [11], Rao [13]-[20]). The question that was asked was the following:

Does there exist a (symmetric) matrix M such that the stationary points of $(y-X\beta)' M(y-X\beta)$, i.e., solutions of the equation $X' MX\hat{\beta} = X' My$, yield estimators $\hat{\beta}$, which are BLUE of β in the sense that $p'\hat{\beta}$ is BLUE of $p'\beta$, whenever $p'\beta$ is estimable? Moreover, does $f^{-1}(y-X\hat{\beta})' M(y-X\hat{\beta})$ provide a reasonable (MINQUE or BQUE) estimator of σ^2 for some appropriate integer f?

The answer was that M exists and has necessarily the form

$$M = (V+XUX')^{-}, \tag{1.2}$$

an arbitrary g-inverse of $V+XUX'$, where U is a symmetric matrix such that $\operatorname{im}(V+XUX') = \operatorname{im}(X \vdots V)$; f is equal to Rank $(X \vdots V)$ $-$ Rank (X). Also the problem of testing linear hypotheses via this approach has been dealt with in papers by Rao ([12], [17], [20]).

It seems to me that the main idea behind this approach is to enlarge the covariance-matrix V artificially to a matrix W in such a way that the condition $\operatorname{im} X \subseteq \operatorname{im} W$ is met. If thus is done in an appropriate way, a correct result will be obtained. It is the purpose of this paper to pursue this idea in detail.

2. Estimation of the mean value. First of all we define the concept of a Unified Least Squares-Matrix (ULS-Matrix).

2.1. Definition. Let the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 V$ be given. A symmetric n.n.d. $n\times n$-matrix W is called *ULS-matrix* with respect to this model if

$$\operatorname{im}(VW^{-}X) \subseteq \operatorname{im} X \subseteq \operatorname{im}(W) \tag{2.1}$$

for some g-inverse W^{-} of W.

2.2. Remark. If, moreover, $\operatorname{im}(V) \subseteq \operatorname{im}(W)$, then (2.1) implies that $\operatorname{im}(V\tilde{W}X) \subseteq \operatorname{im} X$ for any g-inverse $\tilde{W}$ of W.

Proof. $\operatorname{im}(V) \subseteq \operatorname{im} W$ implies $WW^{-}V = W\tilde{W}V = V = W(\tilde{W})'V = V\tilde{W}W$, since $(\tilde{W})'$ is also a g-inverse of W. Therefore $V\tilde{W}V = VW^{-}W\tilde{W}X = VW^{-}X$, since $\operatorname{im}(X) \subseteq \operatorname{im}(W) = \operatorname{im}(W\tilde{W})$, $W\tilde{W}X = X$. Thus indeed $V\tilde{W}X$ and $VW^{-}X$ and a fortiori $\operatorname{im}(VW^{-}X)$ and $\operatorname{im}(V\tilde{W}X)$ coincide, q.e.d. (If $\operatorname{im} X \subseteq \operatorname{im} W$, then $VW^{-}X$ is independent of the choice of W^{-} iff $\operatorname{im} V \subseteq \operatorname{im} W$.)

2.3. Remark. The existence of a ULS-matrix can be seen as follows: Let $W = V+XX'$. Since W is the sum of two n.n.d. matrices it is n.n.d. and $\operatorname{im}(W) = \operatorname{im}(X)+\operatorname{im}(V) = \operatorname{im}(X \vdots V)$. Moreover, from $(XX'+V)[I-W^{-}(XX'+V)] = 0$ or from $(I-W^{-}W)X = 0$, $(I-W^{-}W)V = 0$, it follows that

$$X'[I-W^-(XX'+V)] = 0, \qquad V[I-W^-(XX'+V)] = 0, \tag{2.2}$$

$$V-VW^- V = VW^- XX', \qquad VW^- X = X-XW^- XX'. \tag{2.3}$$

This implies $\mathrm{im}(VW^- X) \subseteq \mathrm{im}\, X$ and $\mathrm{im}(V-VW^- V) \subseteq \mathrm{im}\, X$.

The second condition will play an important role in this paper, too.

2.4. THEOREM. *Let W be a symmetric n.n.d. matrix such that $im(X) \subseteq im\, W$. Then*

$$Gy = X(X'(W^-)'X)^- X'(W^-)'y \tag{2.4}$$

is BLUE of $X\beta$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 W$. Moreover, Gy is BLUE of $X\beta$ also in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$ if and only if W is a ULS-matrix with respect to W^-, i.e., if $\mathrm{im}(VW^- X) \subseteq \mathrm{im}(X)$.

The proof of this theorem was given in Drygas [4], theorem 2.4. Theorem 2.5 was proved in the same paper.

2.5. THEOREM. *Let $\mathrm{im}\, X$, $\mathrm{im}\, V \subseteq \mathrm{im}\, W$ and let W^- be an arbitrary g-inverse of W. Let, moreover, Gy be a BLUE of $\mathrm{E}y$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = W$. Then Gy is BLUE of $\mathrm{E}y$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = V$ independent of the choice of Gy if and only if $\mathrm{im}(VW^- X) \subseteq \mathrm{im}(X)$.*

The following theorem characterizes the BLUE by a single equation.

2.6. THEOREM. (a) *Let $GW = X(X'W^- X)^- X'$ and $\mathrm{im}\, X$, $\mathrm{im}(V) \subseteq \mathrm{im}\, W$, W n.n.d. and symmetric. Then Gy is BLUE of $\mathrm{E}y$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$ if $\mathrm{im}(VW^- X) \subseteq \mathrm{im}(X)$.*

(b) *If, moreover, $\mathrm{im}(X \vdots V) = \mathrm{im}(W)$, then Gy is BLUE of $X\beta$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$ if and only if $GW = X(X'W^- X)^- X'$.*

Proof. (a) Let $y = X\beta = Wa$, then

$$\begin{aligned} Gy &= X(X'W^- X)^- X'a = X(X'W^- X')^- X'W^- Wa \\ &= X(X'W^- X)^- X'W^- X\beta = X\beta. \end{aligned} \tag{2.5}$$

If $y = Va = Wb$; $X'a = 0$, then

$$\begin{aligned} Gy &= X(X'W^- X)^- X'b = X(X'W^- X)'X'W^- Wb \\ &= X(X'W^- X)^- X'W^- Va = X(X'W^- X)^- T'X'a = 0, \end{aligned} \tag{2.6}$$

since $V(W^-)'X = XT$ for some T in view of $\mathrm{im}\big(V(W^-)'X\big) \subseteq \mathrm{im}\, X$. *Thus* Gy is BLUE of $\mathrm{E}y = X\beta$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$, $\sigma > 0$.

(b) If $\mathrm{im}(W) = \mathrm{im}(X \vdots V)$, then the BLUE is unique on $\mathrm{im}(W)$. Hence it follows that if $G_1 y$ is a BLUE, then $G_1 W = GW$ for any other BLUE Gy. Now $Gy = X(X'W^- X)X'W^- y$ is a BLUE. Therefore $G_1 W = GW = X(X'W^- X)^- X'W^- W = X(X'W^- X)^- X'$, q.e.d.

When enlarging V one could also think of enlarging V to a matrix W such that $\operatorname{im}(V) \subseteq \operatorname{im}(W)$. This case has already dealt with in Drygas [2], p. 50.

2.7. THEOREM. *Let W be n.n.d. and symmetric, $\operatorname{im} V \subseteq W$, and let Gy be a BLUE of $\mathrm{E}y = X\beta$ in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 W$. Necessary and sufficient that any such BLUE is as well BLUE in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 V$ is the condition*

$$VW^-(\operatorname{im} X \cap \operatorname{im} W) \subseteq \operatorname{im}(X).$$

Proof. The zero set of all BLUE's in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 W$ is $WX'^{-1}(0)$. Thus $WX'^{-1}(0) \supseteq VX'^{-1}(0)$ or, equivalently,

$$V^{-1}(\operatorname{im} X) \supseteq W^{-1}(\operatorname{im} X)$$

is the necessary and sufficient condition for any BLUE Gy in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 W$ to be as well BLUE of $\mathrm{E}y = X\beta$ in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 V$. This is again equivalent to

$$VW^{-1}(\operatorname{im} X) \subseteq \operatorname{im} X.$$

If $a \in W^-(\operatorname{im} X \cap \operatorname{im} W)$, then $a = W^- b$, $b \in \operatorname{im} X \cap \operatorname{im} W$ and $Wa = WW^- b = b \in \operatorname{im} X$. Thus

$$W^-(\operatorname{im} X \cap \operatorname{im} W) \subseteq W^{-1}(\operatorname{im} X).$$

On the other hand, if $a \in W^{-1}(\operatorname{im} X)$, then

$$Va = VW^- Wa \in \operatorname{im} VW^-(\operatorname{im} X \cap \operatorname{im} W).$$

3. Estimation of the variance. If W is such that $\operatorname{im}(X) \subseteq \operatorname{im}(W)$ and W is ULS with respect to some g-inverse W^- of W, then in the last section it has been shown that a BLUE of $\mathrm{E}y = X\beta$ in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 V$ can be computed via an appropriate Aitken formula for the BLUE in the model $\mathrm{E}y = X\beta$, $\operatorname{Cov} y = \sigma^2 W$. If such a BLUE Gy is computed, the question may arise whether

$$f^{-1}(y-Gy)' W^-(y-Gy) = f^{-1} y'(I-G)' W^-(I-G)y \tag{3.1}$$

is BQUE (Best Quadratic Unbiased Estimator) of σ^2 for some appropriate integer f. If $y'Ay = \hat{\sigma}^2$, then in the case of quasi-normally distributed y

$$\mathrm{E}(y'Ay) = \beta' X' AX\beta + \sigma^2 \operatorname{tr}(AV) \tag{3.2}$$

$$\operatorname{Var}(y'Ay) = 2\sigma^4 \operatorname{tr}(AVAV) + 4\sigma^2 \operatorname{tr}(X\beta\beta' X' AVA). \tag{3.3}$$

Therefore $\hat{\sigma}^2 = y'Ay$ is an unbiased estimator of σ^2 iff $X'AX = 0$, $\operatorname{tr}(AV) = 1$. In this case also

$$\operatorname{Var}(y'Ay) = 2\operatorname{tr}\big((X\beta\beta' X' + \sigma^2 V)A(X\beta\beta' X' + \sigma^2 V)A\big). \tag{3.4}$$

We introduce

$$f = \operatorname{Rank}(X \vdots V) - \operatorname{Rank}(X). \tag{3.5}$$

3.1. THEOREM. *Let the linear model* $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = \sigma^2 V$ *be given, and let* y *be quasi-normally distributed. Then, if* Gy *is a BLUE of* $\mathrm{E}y$,

$$f^{-1}\, y'(I-G)'\, V^{-}(I-G)\, y \tag{3.6}$$

is a BQUE of σ^2.

$y'Ay$ *is a BQUE of* $f\sigma^2$ *iff one of the following equivalent conditions are met*:

$$X'AX = 0,\ VAX = 0,\ VAV = (I-G)V, \tag{3.7}$$

$$(XX'+V)A(XX'+V) = (I-G)V, \tag{3.8}$$

$$(XX'+V)AX = 0,\ VAVAV = VAV,\ \mathrm{tr}(AV) = f, \tag{3.9}$$

$$\frac{1}{\sigma^2} y'Ay \sim \chi_f^2 \text{ if } y \text{ is normally distributed.} \tag{3.10}$$

(*See, e.g. Seely* [21], *Graybill and Wortham* [5]).

Proof. (a) $y'By$ is an unbiased estimator of zero iff $X'BX = 0$, $\mathrm{tr}(BV) = 0$. Therefore the estimator given by (3.6) is Best Quadratic Unbiased iff

$$\mathrm{tr}\big((X\beta\beta'X'+\sigma^2 V)A(X\beta\beta'X'+\sigma^2 V)B\big) = 0 \ \forall B: \quad X'BX = 0,\ \mathrm{tr}(BV) = 0, \tag{3.11}$$

where $A = f^{-1}(I-G)'V^{-}(I-G)$. Evidently, $\mathrm{tr}(AV) = 1$, $X'AX = 0$. Now $(X\beta\beta'X'+\sigma^2 V)A(X\beta\beta'X'+\sigma^2 V) = \sigma^4 VAV = \sigma^4 f^{-1} V(I-G)'V^{-}(I-G)V$ $= \sigma^4 f^{-1}(I-G)VV^{-}V(I-G)' = \sigma^4 f^{-1}(I-G)V(I-G)' = \sigma^4 f^{-1}(I-G)V$.

Thus (3.11) is equivalent to

$$\mathrm{tr}\big((I-G)VB\big) = \mathrm{tr}(VB) - \mathrm{tr}(GVB) = 0. \tag{3.12}$$

Since $\mathrm{tr}(VB) = \mathrm{tr}(BV) = 0$, only $\mathrm{tr}(GVB) = 0$ if $X'BX = 0$, $\mathrm{tr}(BV) = 0$ has to be shown. $X'BX = 0$ implies $GVBX = 0$ or $X'BGV = 0$. This again implies $GVBGV = 0$ and

$$0 = \mathrm{tr}(V^{-}GVBGV) = \mathrm{tr}(GVV^{-}GVB) = \mathrm{tr}(GVV^{-}VG'B) = \mathrm{tr}(GVG'B) = \mathrm{tr}(GVB). \tag{3.13}$$

(b) Since $(X\beta\beta'X'+\sigma^2 V)A(X\beta\beta'X'+\sigma^2 V)$ is unique, it follows that $y'Ay$ is BQUE of $f\sigma^2$ iff

$$X'AX = 0, \quad \mathrm{tr}(AV) = f, \quad (X\beta\beta'X'+\sigma^2 V)A(X\beta\beta'X'+\sigma^2 V) = \sigma^4(I-G)V \tag{3.14}$$

for all β, σ. Letting $\sigma = 1$, it follows that

$$(X\beta\beta'X'+V)A(X\beta\beta'X'+V) = (I-G)V. \tag{3.15}$$

Let $W_\beta = (X\beta\beta' X' + V)$; then $X\beta \in \text{im}(W_\beta)$ and, therefore,

(3.16) $$(X\beta\beta' X' + V) AX\beta = W_\beta AW_\beta W_\beta^- X\beta = (I-G) VW_\beta^- X\beta = 0,$$

since $(VW_\beta^- X\beta) = \lambda X\beta \in \text{im}\, X$ (ULS-property) and $(I-G)X = 0$.

From (3.16) we again get

(3.17) $$\beta' X' AX\beta = 0, \qquad VAX\beta = 0,$$

i.e. $X'AX = 0$, $VAX = 0$. Using this, (3.14) implies $VAV = (I-G)V$,

(3.18) $$\text{tr}(AV) = \text{tr}(AVV^+ V) = \text{tr}(VAVV^+) = \text{tr}((I-G)VV^+) = f.$$

This follows because $(I-G)VV^+$ is an idempotent matrix vanishing on $\text{im}(X) \cap \text{im}(V) + V^{-1}(0)$ and being the identity on $V(X'^{-1}(0))$. Thus

$$\begin{aligned}\text{tr}((I-G)VV^+) &= \text{Rank}((I-G)VV^+) = \dim V(X'^{-1}(0)) \\ &= \dim(\text{im}(V)) - \dim(\text{im}(V) \cap \text{im}(X)) \\ &= \text{Rank}(V) - [\text{Rank}(V) + \text{Rank}(X) - \text{Rank}(X \vdots V)] \\ &= \text{Rank}(X \vdots V) - \text{Rank}(X) = f.\end{aligned}$$

Therefore it is proved that $y'Ay$ is BQUE of $f\sigma^2$ if *(3.7) is met.*

(b′) Clearly (3.7) implies (3.8). On the other hand, if we have $W = V + XX'$, then $\text{im}(VW^- X) \subseteq \text{im}(X) \subseteq \text{im}(W)$ (ULS-property) and (3.8) implies

(3.19) $$WAX = WAWW^- X = (I-G)VW^- X = 0$$

and from this $X'AX = 0$, $VAX = 0$, $VAV = (I-G)V$, i.e. (3.7), is obtained.

(c) Since

$$VAVAV = (I-G)VAV = (I-G)(I-G)V = (I-G)V(I-G)' = (I-G)V = VAV$$

if (3.8) or (3.9) holds, it is clear that (3.7) or (3.8) implies (3.9). The converse is more complicated. The proof is due to J. Müller (See Müller [12], p. 23). Let $k = \text{Rank}(V)$ and $V = RR'$, where R is a $n \times k$-matrix of Rank k. We define C by $C = R^+ = (R'R)^{-1}R'$, consequently $CR = R'C' = I_k$. Since $VAV = RR'ARR' = VAVAV = RR'ARR'ARR'$, it follows that $R'AR = (R'AR)^2$. Hence $R'AR = P$ is a projection and $\text{Rank}(P) = \text{tr}(P) = \text{tr}(AV) = f$. Let, moreover, $Q = C(GV)C' = CGRR'C' = CGR$. Q is also a projection: $Q^2 = CGRCGR = CG^2R = CGR$, since $G^2RR' = G^2V = GVG' = GV = GRR'$. Finally, $PQ = R'ARCGR = R'AGR$ since $RR'AGR = VAGR = 0$ in view of $\text{im}(GR) = \text{im}(GRR') = \text{im}(GV) = \text{im}(X) \cap \text{im}\, V \subseteq \text{im}(X)$. Thus $\text{im}\, Q \subseteq (\text{im}(P))^\perp = P^{-1}(0)$.

On the other hand, $\text{Rank}(Q) = \text{Rank}(CGR) = \text{Rank}(CGRR') = \text{Rank}(CGV) = \text{Rank}(CVG') = \text{Rank}(R'G') = \text{Rank}(GR) = \text{Rank}(GV) = \dim \times (\text{im}(X) \cap \text{im}(V)) = \text{rank}(V) - f = k - \text{rank}(P)$. Thus $\text{im}(Q) = P^{-1}(0)$. But

Q is also symmetric: $Q' = R'G'C' = CRR'G'C' = CVG'C' = CGVC' = CGRR'C' = CGR$. For this reason $P = I_k - Q$, i.e., $R'AR = I - R'G'C'$ and finally $VAV = V - VG'C'R' = V - VG' = V - GV = (I-G)V$.

(d) The equivalence of (3.10) and (3.11) for the of normally distributed observation y follows immediately from Corollary (2.11.1) in Srivastava and Khatri [22], p. 64.

Since $WAW = WA'W = (I-G)V$, A must not be symmetric. From this it follows that, e.g., $y'W^-(I-G)y$ is BQUE of $f\sigma^2$. This estimator is not necessarily invariant.

The preceding theorem shows that

$$f^{-1}y'(I-G)'W^-(I-G)y \tag{3.20}$$

is BQUE of σ^2 if Gy is a BLUE of Ey in the model E$y = X\beta$, Cov $y = \sigma^2 V$ and W is the ULS-matrix $V + XX'$. Indeed, $(I-G)W = (I-G)V = V(I-G)' = W(I-G)'$ and therefore

$$W(I-G)'W^-(I-G)W = (I-G)WW^-W(I-G)' = (I-G)W(I-G)' = (I-G)V(I-G)' = (I-G)V. \tag{3.21}$$

Now the question arises for which ULS-matrices W the formula (3.20) leads to a BQUE of σ^2. The answer is given by Theorem 3.2.

3.2. THEOREM. (a) *If W is a ULS-matrix with respect to W^-, then*

$$y'(I-G)'W^-(I-G)y \tag{3.22}$$

is BQUE of $f\sigma^2$ iff $\operatorname{im}(V - VW^-V) \subseteq \operatorname{im} X$.

(b) $\operatorname{im}(V - VW^-V) \subseteq \operatorname{im} X$ *holds if and only if $V(W^-)'$ is the identity on* $VX'^{-1}(0)$.

3.3. Remark. Theorem 3.2 generalizes a theorem by Kruskal [6] obtained for $W = I$, V regular. Note that $f = 0$ iff $\operatorname{im} V \subseteq \operatorname{im} X$. In this case the conditions $\operatorname{im}(VW^-X)$, $\operatorname{im}(V - VW^-V) \subseteq \operatorname{im}(X)$ are automatically met. Evidently VW^-V is independent of the choice of the g-inverse W^- of W if $\operatorname{im}(V) \subseteq \operatorname{im}(W)$.

Proof of Theorem 3.2. Clearly $\operatorname{im}(V - VW^-V) \subseteq \operatorname{im} X$ is equivalent to $X'^{-1}(0) \subseteq (V - V(W^-)'V)^{-1}(0)$. This, however, means that $Vz = V(W^-)'z$ for all $z \in X'^{-1}(0)$. Therefore only part (a) of the theorem has to be proved.

Since $(I-G)'W^-(I-G)X = 0$, (3.22) is BQUE of $f\sigma^2$ iff $V(I-G)' \times W^-(I-G)V = (I-G)V$. But $V(I-G)' = (I-G)V$ and we get the equation

$$(I-G)(V - VW^-(I-G)V) = 0. \tag{3.23}$$

$(I-G)z$ is unique if $z \in \operatorname{im} V$ and vanishes there iff $z \in \operatorname{im}(V) \cap \operatorname{im} X$. This implies that (3.23) is equivalent to $\operatorname{im}(V - VW^-(I-G)V) \subseteq \operatorname{im} X$. Now $\operatorname{im}(VW^-GV) \subseteq \operatorname{im} X$ since $\operatorname{im}(GV) \subseteq \operatorname{im} X$ and W was ULS with respect to W^-. Therefore the above relation holds iff $\operatorname{im}(V - VW^-V) \subseteq \operatorname{im} X$, Note that the condition of this theorem is just the condition obtained in (2.3) for the ULS-matrix $V+XX'$. Our next aim is, again, to characterize a BQUE of $f\sigma^2$ by a single equation. This equation is $WAW=(I-G)V$ for the ULS-matrix $W=V+XX'$. This characterization will be valid for arbitrary ULS-matrices W if $\operatorname{im}(W)=\operatorname{im}(X\vdots V)$ and $\operatorname{im}(V-VW^-V)\subseteq \operatorname{im}X$.

3.4. THEOREM. (a) *Let* W *be an ULS-matrix such that* $\operatorname{im}(X\vdots V)\subseteq \operatorname{im}W$ *and* $\operatorname{im}(V-VW^-V)\subseteq \operatorname{im}X$. *If* $WAW=(I-G)V$, *then* $y'Ay$ *is BQUE of* $f\sigma^2$.

(b) *If* $\operatorname{im}(V\vdots X)=\operatorname{im}(W)$; $\operatorname{im}(VW^-X)$, $\operatorname{im}(V-VW^-V)\subseteq \operatorname{im}X$, *then* $y'Ay$ *is BQUE of* $f\sigma^2$ *iff* $WAW=(I-G)V$.

(c) *Under the assumptions of* (b) $\operatorname{im}(V-W)\subseteq \operatorname{im}X$.

(d) *Under the assumptions of* (b) *the BLUE's of* $\mathrm{E}y$ *in the models* $\mathrm{E}y=X\beta$, $\operatorname{Cov}y=\sigma^2V$ *and* $\mathrm{E}y=X\beta$, $\operatorname{Cov}y=\sigma^2W$ *coincide.*

(e) *Under the assumptions of* (b) $W=X\Lambda X'+V$ *for some symmetric* Λ.

The assertion (e) of this theorem shows that we finally arrive at the class of matrices considered by Rao and Mitra. It may be noted that if $W=X\Lambda X'+V$ is such that $\operatorname{im}(X\vdots V)=\operatorname{im}W$, then $X'^{-1}(0)$, $V^{-1}(0)\supseteq W^{-1}(0)$ and $W[I-W^-W]=0$ implies $X'[I-W^-W]=0$, $V[I-W^-W]=0$. These equations are equivalent to $V(W^-)'X=X-X\Lambda X'(W^-)'X$ and $V-VW^-X=VW^-X\Lambda X'-X\Lambda X'(W^-)'X\Lambda X'$. Therefore the ULS-property and the property $\operatorname{im}(V-VW^-V)\subseteq \operatorname{im}X$ are fulfilled in this case.

Proof of the theorem. (a) Since $\operatorname{im}(X)\subseteq \operatorname{im}(W)$, $WW^-X=X$ and therefore $WAX=WAWW^-X=(I-G)VW^-X=0$ in view of $\operatorname{im}(VW^-X)\subseteq \operatorname{im}X$, $(I-G)X=0$. Since $X'^{-1}(0)$, $V^{-1}(0)\supseteq W^{-1}(0)$ from this $X'AX=0$, $VAX=0$ is obtained. Finally $WAV=WAWW^-V=(I-G)VW^-V=(I-G)V-(I-G)(V-VW^-V)=(I-G)V$, since $\operatorname{im}(V-VW^-V)\subseteq \operatorname{im}X$, $(I-G)X=0$. By theorem 3.1 $y'Ay$ is BQUE of $f\sigma^2$, since $VAV=VW^-WAV=VW^-(I-G)V=(V-VW^-V)(I-G)V=(I-G)V$.

(b) If $y'Ay$ is BQUE of $f\sigma^2$, then $X'AX=0$, $X'AV=0$, $VAV=(I-G)V$. Since $\operatorname{im}(W)=\operatorname{im}(X\vdots V)$ *from this* $X'AW=0$, $WAX=0$ is obtained. Now let $Wa=X\beta+Vb$. From this we get $AWa=VAVa=(I-G)Vb=(I-G)Wa$. Finally, $WAWa=WaVb=W(I-G)'b$. *In (c) we will show that* $\operatorname{im}(V-W)\subseteq \operatorname{im}X$. *Thus* $(I-G)W=(I-G)V$ and $WAVb=(I-G)V'b=(I-G)Vb=(I-G)(Va+X\beta)=(I-G)Wa=(I-G)Va$.

(c) Let $Wa=Vb+X\beta$, $X'b=0$. There $Va=VW^-Wa=VW^-(Vb+X\beta)=Vb+VW^-X\beta$. Therefore $Va-Wa=VW^-X\beta-X\beta\in\operatorname{im}(X)$ by the ULS-property.

(d) If $\operatorname{im}(VW^-X)\subseteq \operatorname{im}(X)$, then, by Theorem 2.5, $V(X'^{-1}(0))\subseteq W(X'^{-1}(0))$. On the other hand, from $\operatorname{im}(V-W)\subseteq \operatorname{im}X$ we get $X'^{-1}(0)\subseteq (V-W)^{-1}(0)$ and therefore $Va=Wa$ if $X'a=0$. Thus the coincidence of the BLUE's is evident.

(e) If $P=XX^+$ is the projection on $\operatorname{im}(X)$, then from $\operatorname{im}(V-W)\subseteq \operatorname{im}X$ it is obtained that $P(V-W)=(V-W)P=V-W$. This shows $V-W=$

$P(V-W)P = XX^+(V-W)(X^+)'X' = -X\Lambda X'$, $\Lambda = -X^+(V-W)(X^+)'$ and $W = V+X\Lambda X'$, q.e.d.

4. Prediction. Consider the linear model

$$\mathrm{E}\left(\frac{y}{y_*}\right) = \left(\frac{X}{X_*}\right)\beta, \quad \mathrm{Cov}\left(\frac{y}{y_*}\right) = \sigma^2\left(\frac{V}{V_{12}'}+\frac{V_{12}}{V_{22}}\right), \tag{4.1}$$

y is observed, but y_* is unobserved and to be predicted. Consider a linear function $l'y_*$ which is predictable, i.e., $X_*'l \in \mathrm{im}(X')$. Then $a'y$ is the Best Linear Unbiased Predictor (BLUP) of $l'y_*$ if

$$X'a = (X_*)'l, \tag{4.2}$$

$$Va - V_{12}l \in \mathrm{im}(X), \tag{4.3}$$

(Drygas [3], Toutenburg [23], Baksalary and Kala [1]). A solution of (4.2), (4.3) is e.g. given by

$$a = G'(X^+)'(X_*)'l + (I-G)'V^-(V_{12})l, \tag{4.4}$$

where Gy is a BLUE of $X\beta$ in the model $\mathrm{E}y = X\beta$, $\mathrm{Cov}\, y = V$. If we replace V by a matrix W, n.n.d and symmetric, such that $\mathrm{im}(V_{12} - VW^- V_{12}) \subseteq \mathrm{im}(X)$, then in (4.4) V can be replaced by W. If, e.g., $\mathrm{im}(V-W) \subseteq \mathrm{im}(X)$, then $Wa - V_{12}l \in \mathrm{im}\, X$ implies $Va - V_{12}l \in (X)$. Since $\mathrm{im}(V_{12}) \subseteq \mathrm{im}\, V$, $\mathrm{im}(V_{12} - VW^- V_{12}) \subseteq \mathrm{im}(V - VW^- V)$; we arrive at a well-known condition.

If $X_* = TX$, then y_* is predictable and $G_1 y$ is BLUP of y_* iff $G_1 X = X$, $G_1 Vz = V_{12}'z$ if $X'z = 0$. A single-equation characterization is

$$G_1 W = X_*(X'W^- X)^- X)^- X' + V_{21}(I - W^- X(X'W^- X') \quad (V_{21} = V_{12}')$$

4.1. THEOREM. (a) *If* $\mathrm{im}(VW^- X) \subseteq \mathrm{im}(X) \subseteq \mathrm{im}(W)$ *and* $\mathrm{im}((X_*)') \subseteq \mathrm{im}(X')$, *then*

$$G_1 y = X_*(X'(W^-)'X)^- X'(W^-)'y + \\ + V_{12}'W^-\big(I - X(X'(W^-)'X)'X'(W^-)'\big)y \tag{4.5}$$

is BLUP of y_* *iff* $\mathrm{im}(V_{12} - V(W^-)'V_{12}) \subseteq \mathrm{im}\, X$.

(b) *If* $\mathrm{im}(X)$, $\mathrm{im}(V) \subseteq \mathrm{im}(W)$; $\mathrm{im}(VW^- X)$, $\mathrm{im}(V_{12} - VW^- X_{12}) \subseteq \mathrm{im}\, X$ *and*

$$G_1 W = X_*(X'W^- X)^- X' + V_{12}'(I - W^- X(X'W^- X)^- X'), \tag{4.6}$$

then $G_1 y$ *is BLUP of* y_*.

(c) *If* $\mathrm{im}(X \vdots V) = \mathrm{im}(W)$ *and* $\mathrm{im}(VW^- X)$, $\mathrm{im}(V_{12} - VW^- V_{12}) \subseteq \mathrm{im}(X)$, *then* $G_1 y$ *is BLUP of* y^* *iff* $G_1 W$ *is given by formula* (4.6).

For example if $G_1 y = X_*(X^-)Gy + V_{12}'Ay$, $y'Ay$ BQUE of $f\sigma^2$, then $G_1 y$ is BLUP of y_*.

REFERENCES

[1] J. K. Baksalary and R. Kala, *Simple Least Squares Estimation versus Best Linear Unbiased Prediction*, Journ. of Statist. Planning and Inference 5 (1981), p. 147 - 151.

[2] H. Drygas, *Zur Theorie der Gauss - Markoffschen Schätzungen*, Ph. D. thesis, University of Heidelberg, 1968.

[3] — *Estimation and prediction for linear models in general spaces*, Math. Operationsf. und Statistik, 6 (1975), p. 301 - 324.

[4] — *Sufficiency and completeness in the general Gauss-Markov model*, Sankhya, A, 45 (1983).

[5] F. A. Graybill, and A. W. Wortham, *A note on uniformly best unbiased estimators for Variance components*, J. Amer. Statist. Assoc., 51 (1956), p. 266 - 268.

[6] W. Kruskal, *When are Least Squares and Gauss - Markov Estimators identical? A coordinate - free approach*, Ann. math. Statist. 39 (1968), p. 70 - 75.

[7] S. K. Mitra, *Unified Least Squares Approach to Linear Estimation in a general Gause - Markov model*, Siam. Journ. of Appl. Math., 25 (1973), p. 671 - 680.

[8] — and B. Moore, *Gauss - Markov estimation with an incorrect dispersion matrix*, Sankhyā, A, 35 (1973), p. 139 - 152.

[9] S. K. Mitra and C. R. Rao, *Further contributions to the theory of generalized inverse of matrices and its applications*, Sankhyā, A, 33 (1971), p. 289 - 300.

[10] — *Some results on Estimation and Tests, of Linear Hypotheses in the Gauss - Markoff Model*, Sankhyā, A, 30 (1968), pp. 281 - 290.

[11] — *Generalized Inverse of Matrices and its applications*, Wiley, New York 1971.

[12] J. Müller, *Suffizienz und Vollständigkeit in linearen Modellen*, Ph. D. thesis, Gesamthochschule Kassel, 1982.

[13] C. R. Rao, *Linear statistical inference and its applications*, Wiley, New York 1965 and 1973.

[14] — *Unified theory of linear estimation*, Sankhyā, A, 33 (1971), p. 477.

[15] — *Some recent results in linear estimation*, Sankhyā, B, 34 (1972), p. 369 - 378.

[16] — *Unified Theory of Least Squares*, Comm. Statist. 1 (1973), p. 1 - 8.

[17] — *A note on the IPM - method and the Unified theory of linear estimation*, Sankhyā A, 34 (1972), p. 285 - 288.

[18] — *Theory of Estimation of Parameters in the general Gauss - Markoff model* in *A survey of Statistical Design and Linear Models*, J. N. Srivastava (Ed.), North Holland, Amsterdam 1973.

[19] — *A note on the Unified Theory of Least Squares*, Communications in Stat., A, 7 (5) (1978), p. 409 - 411.

[20] — *Least Squares Theory for possible singular models*, The Canad. Journ. of Statist. 6 (1978), p. 19 - 23.

[21] J. Seely, *A complete sufficient statistic for the linear model under normality and a singular covariance - matrix*, Comm. Statist, A, 7 (1978), p. 1465 - 1473.

[22] M. S. Srivastava and C. G. Khatri, *An introduction to multivariate Statistics*, North Holland, New York, Amsterdam (1979).

[23] H. Toutenburg, *Vorhersage in linearen Modellen*, Akademie - Verlag, Berlin 1975.

Gesamthochschule
Universität des Landes Hessen
FB 17, Postf. 101 380
D - 3500 Kassel, BRD

Received on 25. 4. 1983