SUFFICIENCY IN LINEAR MODELS

HILMAR DRYGAS

Gesamthochschule, Kassel, F.R.G.

1. Introduction of the subject

Let $(\Omega, \mathcal{F}, P)$ be a probability space, ε a random $n \times 1$ -vector such that

$$E\varepsilon = 0, \quad \text{Cov } \varepsilon = E\varepsilon\varepsilon' = V, \quad (1.1)$$

where V is a symmetric n.n.d. $n \times n$ -matrix. Let, moreover, X be a fixed $n \times k$ -matrix, called the regression matrix. We consider transformations of the kind

$$g_{\beta, \sigma}(a) = X\beta + \sigma a, \quad a \in \mathbb{R}^n. \quad (1.2)$$

If β runs over $\mathbb{R}^k$ and σ runs over $\mathbb{R}^+$, then

$$G = \{g_{\beta, \sigma}: \beta \in \mathbb{R}^k, \sigma \in \mathbb{R}^+\} \quad (1.3)$$

forms a group where the group operation is defined by

$$\begin{aligned} g_{\beta, \sigma} g_{\beta', \sigma'}(a) &= g_{\beta, \sigma}(g_{\beta', \sigma'}(a)) = X\beta + \sigma(X\beta' + \sigma'a) \\ &= X(\beta + \sigma\beta') + \sigma\sigma'a = g_{\beta + \sigma\beta', \sigma\sigma'}(a). \end{aligned} \quad (1.4)$$

Therefore $e = g_{0,1}$ is the unit-element in G , while

$$\begin{aligned} (g_{\beta, \sigma})^{-1} &= g_{-\beta\sigma^{-1}, \sigma^{-1}}, \\ G_1 &= \{g_{\beta, 1}: \beta \in \mathbb{R}^k\} \end{aligned} \quad (1.5)$$

is evidently a subgroup of G and

$$Y = \{g\varepsilon: g \in G\} \quad \text{and} \quad Y_1 = \{g\varepsilon: g \in G_1\} \quad (1.6)$$

form transformation models which we will also denote by $M(\text{im}(X), \{\sigma^2 V\})$ and $M(\text{im } X, V)$, respectively. By $\text{im}(A)$ we will in the sequel denote the column-space of a matrix A .

Let us now for simplicity consider the group G_1 and let us, moreover, assume that V is regular and ε is normally distributed. Then the density of $y \in Y_1$ at the point ε is equal to

$$f(\xi, \beta) = (2\pi)^{-n/2} (\det(V))^{-1/2} \exp(-\frac{1}{2}(\xi - X\beta)' V^{-1}(\xi - X\beta)). \quad (1.7)$$

Consider the normal equations

$$X' V^{-1} X \hat{\beta} = X' V^{-1} \xi. \quad (1.8)$$

Since $(X' V^{-1} X)^{-1}(0) = X^{-1}(0)$ and therefore $\text{im}(X' V^{-1} X) = \text{im}(X')$ these normal equations will always possess a solution. If $\hat{\beta}$ is any solution, then it follows from the normal equations that $X \hat{\beta}$ is unique and

$$(\xi - X\beta)' V^{-1} (\xi - X\beta) = (\xi - X\hat{\beta})' V^{-1} (\xi - X\hat{\beta}) + (\hat{\beta} - \beta)' X' V^{-1} X (\hat{\beta} - \beta) \quad (1.9)$$

If we let

$$h(\xi) = (2\pi)^{-n/2} (\det(V))^{-1/2} \exp(-\frac{1}{2}(\xi - X\hat{\beta})' V^{-1} (\xi - X\hat{\beta})), \quad (1.10)$$

$$g_{\beta}(X\hat{\beta}) = \exp(-\frac{1}{2}(X\beta - X\hat{\beta})' V^{-1} (X\beta - X\hat{\beta})) \quad (1.11)$$

it follows that

$$f(\xi, \beta) = h(\xi) g_{\beta}(X\hat{\beta}). \quad (1.12)$$

From the Neyman-Fisher factorization theorem we get $X\hat{\beta}$ is a sufficient statistic. In the same way it can be shown that for the model Y

$$T(y) = (X\hat{\beta}, (y - X\hat{\beta})' V^{-1} (y - X\hat{\beta})) \quad (1.13)$$

is a sufficient statistic. Moreover, it also well-known that this statistic is also complete, i.e.,

$$E_{\beta, \sigma} \varphi(T(y)) = 0, \quad \forall \beta, \sigma \Rightarrow \varphi(\xi) = 0 \text{ a.s.} \quad (1.14)$$

The fact that $X\hat{\beta}$ and $T(y)$ are sufficient and/or complete heavily depends on the normality assumption. Linear model theory, however, can be developed along two lines: without and with normality assumption, e.g., the concept of best unbiased estimation in normal theory corresponds to the concept of best linear unbiased estimation in the non-normal case. We will briefly discuss how this can be extended to the concepts of sufficiency and completeness. More details can be found in [4].

2. Linear sufficiency and linear completeness

2.1. DEFINITION. Let $Y_1 = M(\text{im } X, V)$ be given, G a $n \times n$ -matrix. Gy is called *BLUE* (Best Linear Unbiased Estimator) of Ey in this model if

- (i) $E(Gy) = Ey \quad \forall y \in Y_1$,
- (ii) $\text{Cov}(Gy) \leq \text{Cov}(\hat{G}y)$, $\forall y \in Y_1$, and all $\hat{G}$ meeting (i).

2.2. THEOREM. Gy is BLUE of Ey in the given model if and only if

- (i) $GX = X$,
- (ii) $GVa = 0$, $\forall a \in X'^{-1}(0) = (\text{im}(X))^{\perp}$, i.e., if and only if $G = P_{\text{im } X|V(X'^{-1}(0))}$.

Proof. See, e.g., [6]. ■

2.3. DEFINITION. A linear statistic Ay is said to be *linearly sufficient* for Y_1 if there is linear transformation B such that BAy is BLUE of Ey in Y_1 .

This concept was essentially defined by Baksalary and Kala ([1]) from Poznań who, however, did not use the term "linearly sufficient". Before we can characterize the set of all linearly sufficient statistics we need the concept of a ULS-matrix.

2.4. DEFINITION. In the model Y_1 a $n \times n$ -matrix W , n.n.d. and symmetric, is called *ULS (Unified Least Squares)* if

$$\text{im}(VW^-X) \subseteq \text{im} X \subseteq \text{im} W \quad (2.1)$$

for some g -inverse W^- of W ($WW^-W = W$).

EXAMPLE OF A ULS-MATRIX. $W = V + XX'$ (Rao, Mitra).

Remark. If $\text{im} X \subseteq \text{im} W$, then $\text{im}(VW^-X)$ is independent of the choice of the g -inverse if and only if $\text{im}(V) \subseteq \text{im}(W)$. This is a result due to Baksalary and Kala (to appear in *Linear and Multilinear Algebra*).

2.5. THEOREM. Let the model $Y_1 = G_1(\varepsilon)$ be given and let W be an ULS-matrix with respect to this model such that $\text{im}(V) \subseteq \text{im}(W)$.

(a) If

$$\text{im}(X) \subseteq \text{im}(WA') \quad (2.2)$$

then the equation $BAX = X$ has a solution and $BHx = BHAy$ is BLUE of $Ey = X\beta$ in Y_1 if Hx denotes a BLUE of $AX\beta = Ez$ in the model $Ez = AEy$, $\text{Cov } z = AWA'$, i.e., $H = P_{\text{im} AX | (AWA')(AX)^{-1}(0)}$.

(b) If, moreover, $\text{im}(W) = \text{im}(X:V)$ then Ay is a LS statistic if and only if (2.2) is met.

Proof. (a) We will simplify the proof by taking $W = V + cXX'$, where $c \geq 0$ is such that $\text{im}(X) \subseteq \text{im}(W)$.

The general proof is given in [4].

$$\begin{aligned} X = BAX &\leftrightarrow X' = X'A'B' \leftrightarrow \text{im}(X) \subseteq \text{im}(X'A') \\ &\leftrightarrow (AX)^{-1}(0) \subseteq X'^{-1}(0). \end{aligned}$$

$$AXa = 0, Xa = WA'b \text{ for some } b \Rightarrow AWA'b = 0 \Rightarrow WA'b = Xa = 0.$$

(b) Consider Hx such that $HAX = AX$, $AWA'b = 0$ if $X'A'b = 0$. Then it follows that $HAWA'b = HAVA'b = 0$, if $X'A'b = 0$. Evidently $BHAX = BAX = X$. Now consider $BHAVA$, $X'a = 0$. Then $Va = Wa$ and $AVa = AWa = AWA'w$ for some w , implying $AW(a - A'w) = 0$. Since

$$\text{im}(X) \subseteq \text{im}(WA') \leftrightarrow (AW)^{-1}(0) \subseteq X'^{-1}(0),$$

it follows that $X'a - X'A'w = 0$ and $X'A'w = 0$. Thus finally $HAVa = HAWA'w = HAVA'w = 0$, since $X'A'w = 0$.

(c) Let $X^\perp$ be a matrix such that $\text{im}(X^\perp) = X'^{-1}(0) = (\text{im } X)^\perp$.

$$Ay \text{ LS} \leftrightarrow \exists L, \quad LAX - X = LAVX^\perp = 0.$$

We prove

$$(AW)^{-1}(0) \subseteq X'^{-1}(0).$$

Let $AWt = 0$, $Wt = Xk + VX^\perp h$, since $\text{im}(W) = \text{im}(X:V)$. $LAWt = Xk = 0$, $Wt = VX^\perp h$. From (2.1) it follows that $\text{im}(VX^\perp) \subseteq \text{im}(WX^\perp)$. Thus $Wt = Wt_1$, $X't_1 = 0 \Rightarrow W(t - t_1) = 0 \Rightarrow X't - X't_1 = 0 = X't$, since $W^{-1}(0) \subseteq X'^{-1}(0)$. Thus $X't = 0$. ■

2.6. DEFINITION. Let $M(\text{im } X, V)$ be given, $z = Ay$ is called *LMS* (linearly minimal sufficient) if it is linearly sufficient and for any transformation $z_1 = By$ which is linearly sufficient there exists a linear transformation C such that $Ay = CBy$ almost surely.

2.7. THEOREM. In the model $Ey = X\beta$, $\text{Cov } y = V$ let W be ULS such that $\text{im}(W) = \text{im}(X:V)$. $z = Ay$ is linearly sufficient if and only if

$$\text{im}(WA') = \text{im}(X).$$

Proof. 1) Let $\text{im}(WA') = \text{im}(X)$ and $\text{im } X \subseteq \text{im}(WB') \Rightarrow \text{im}(WA') \subseteq \text{im}(WB') \Rightarrow W'B'C' = WA' \Rightarrow AW = CBW$. Since $y \in \text{im}(W)$ almost surely, $Ay = CBy$ almost surely.

2) The converse is more complicated, it can be found in [4]. ■

2.8. DEFINITION. In the model $Ey = X\beta$, $\text{Cov } y = V$ a linear statistic $z = Ay$ is said to be *LC* (linearly complete) if for all linear transformations B from

$$EBAy = 0, \quad \forall y \in Y_1,$$

it follows that

$$BAy = 0 \text{ a.s.}, \quad \forall y \in Y_1.$$

2.9. THEOREM. In the given model $z = Ay$ is linearly complete if and only if $\text{im}(AV) \subseteq \text{im}(AX)$.

2.10. THEOREM. (a) Ay is LS $\leftrightarrow \text{im}(X) \subseteq \text{im}(WA')$

(b) Ay is LMS $\leftrightarrow \text{im } X = \text{im}(WA')$

(c) Ay is LC $\leftrightarrow \text{im}(AV) \subseteq \text{im}(AX)$

(d) Ay is LMS $\leftrightarrow Ay$ is LS and LC.

2.11. THEOREM. $z = Ay$ is LS in the given model if and only if one of the following conditions is met:

- (i) $A^{-1}(0) \cap \text{im}(W) \subseteq (X'^{-1}(0))$,
- (ii) $\text{im } X \subseteq \text{im}(WA)$,
- (iii) $\exists B: BAX = X, BAVA' = VA'$,
- (iv) $\exists B: BAX = X, BAVA'B' = BAV$,
- (v) $\exists B: BAX = X, (BAV)' = BAV$.

The proof is given in [8].

2.12. THEOREM. (a) *If ε is normally distributed then $z = Ay$ is sufficient if z is linearly sufficient.*

(b) *If ε is normally distributed then $z = Ay$ is complete if and only if z is linearly complete.*

Proof. (a) By Boullion and Odell ([3], p. 54) a version of the conditional distribution of y given Ay , can be obtained by $N((I - BA)X\beta + BAy, (I - BA)V)$ where B is any solution of the equation $BAVA' = VA'$. By the preceding theorem there is a solution of this equation such that $(I - BA)X = 0$, if $z = Ay$ is linearly sufficient.

The converse of part (a) of the above theorem is also true. But the proof is more complicated if V is singular (see [8]).

3. Quadratic concepts

Now let us turn over to the model Y . Define $f = \text{Rank}(V: X) - \text{Rank}(X)$ as the degrees of freedom involved in this model. A quadratic function $y' Ay$ is said to be a *BQUE* (best unbiased quadratic estimator) of $f\sigma^2$ if it unbiased, i.e., $X'AX = 0$, $\text{tr}(AV) = f$ and it has smallest variance in the class of all quadratic unbiased estimators of $f\sigma^2$. Since quasi-normality is assumed the expression to be minimized is

$$\text{tr}(A(\sigma^2 V + X\beta\beta' X') A(\sigma^2 V + X\beta\beta' X')). \quad (3.1)$$

3.1. THEOREM. *Let Gy be a BLUE of Ey , then $y' Ay$ is BQUE of σ^2 if and only if $WAW = (I - G)V$ or equivalently $VAX = 0$, $X'AX = 0$, $VAV = (I - G)V$, where W again is ULS, $\text{im}(W) = \text{im}(X: V)$ and Gy is a BLUE of Ey in $M(\text{im } X, V)$.*

Proof. See [8]. ■

In the model $M(0, \{\sigma^2 V\})$ we call a statistic $z = Ay$ *quadratically sufficient* if there exists a matrix Q such that $z' Qz$ is BQUE of σ^2 . It can be shown that this is the case iff $\text{im}(V) \cap A^{-1}(0) = \{0\}$ ([8]). You may also ask whether some quadratic form $y' Ty$ is quadratically sufficient in $N(0, \sigma^2 V)$. This is the case if and only if $T = \alpha V^-$, $\alpha \in \mathbb{R}$, V^- a generalized inverse of V .

The next step is to consider the model $M(\text{im } X, \{\sigma^2 V\})$ and to ask whether $z = Ay$ is quadratically sufficient, i.e., whether there exist Q, L such

that $z'Qz$ is BQUE of $f\sigma^2$ and LAy is BLUE of $X\beta$. It is clear that such a sufficient statistic can only exist in very special cases. The condition for quadratic sufficiency and sufficiency is that $(AW)^{-1}(0) \subseteq W^{-1}(0)$ ([8]). If Ay is sufficient and complete, then $\text{im } V \subseteq \text{im } X$ and σ^2 is not estimable.

Quadratic sufficiency of $(Ay, y'Ty)$ implies ordinary sufficiency. The converse is not correct without additional assumptions. The following theorem was as well proved by Müller ([8]).

3.2. THEOREM. *$(Ay, y'Ty)$ is quadratically sufficient and linearly complete if and only if*

$$(i) \quad A^{-1}(0) \cap \text{im } W = VX'^{-1}(0) \leftrightarrow \text{im}(X) = \text{im}(WA');$$

$$(ii) \quad \text{There exists } \alpha \text{ such that } \alpha X' TVX^\perp = 0, \alpha V TVX^\perp = VX^\perp,$$

where $\text{im}(X^\perp) = (\text{im } X)^\perp = X'^{-1}(0)$.

J. K. Baksalary ([2]) has extended the concepts of linear sufficiency, minimal linear sufficiency and linear completeness to the case of presence of nuisance parameters.

References

- [1] J. K. Baksalary and R. Kala, *Linear transformations preserving the best linear unbiased estimator in a general linear model*, Ann. Statist. (1981), to appear.
- [2] J. K. Baksalary, *A study of the equivalence between a Gauss-Markoff model and its Augmentation by nuisance parameters*, Math. Operationsforsch. Statist. Ser. Statist., to appear.
- [3] Th. Boullin and P. L. Odell, *Generalized Inverse Matrices*, Wiley Interscience, New York 1971.
- [4] H. Drygas, *Sufficiency and completeness in the general Gauss-Markov model*, Sankhyā Ser. A, to appear.
- [5] H. Drygas, *On the generalization of least squares theory*, paper presented at the 14th European meeting of statisticians, Wrocław 1981.
- [6] H. Drygas, *The Coordinate-Free Approach to Gauss-Markov Estimation*, Springer-Verlag, Berlin 1970.
- [7] S. K. Mitra, *Unified least squares approach to linear estimation in a general Gauss-Markov model*, SIAM J. Appl. Math., **25** (1973), 671-680.
- [8] J. Müller, *Suffiziente Statistiken in linearen Modellen*, Dissertation Gesamthochschule Kassel (forthcoming).
- [9] C. R. Rao, *Linear Statistical Inference and its Applications*, Wiley, New York 1973.
- [10] C. R. Rao, *On a unified theory of estimation in linear models*, Lecture Notes, Purdue University, Lafayette, Indiana, 1973.
- [11] C. R. Rao, *Unified theory of least squares*, Comm. Statist. **1** (1973), 1-8.
- [12] C. R. Rao and S. K. Mitra, *Generalized Inverses of Matrices and Its Applications*, Wiley, New York 1971.
- [13] J. Seely, *A complete sufficient statistic for the linear model under normality and a singular covariance-matrix*, Comm. Statist., to appear.

*Presented to the semester
Sequential Methods in Statistics
September 7-December 11, 1981*