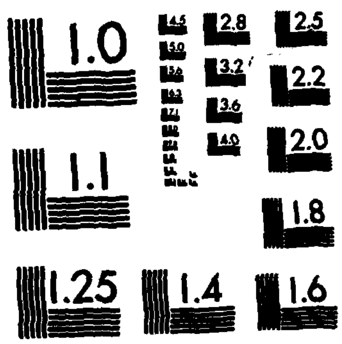
AD-A125 152 LINEAR SUFFICIENCY AND SOME APPLICATIONS IN MULTILINEAR ESTIMATION(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS H DRYGAS OCT 82 TR-82-30 1/1

UNCLASSIFIED AFOSR-TR-83-0007 F49620-82-K-0001 F/G 12/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Center for Multivariate Analysis

University of Pittsburgh

LINEAR SUFFICIENCY AND SOME APPLICATIONS IN MULTILINEAR ESTIMATION[1]

Hilmar Drygas[2]

Center for Multivariate Analysis
University of Pittsburgh

October 1982

Technical Report No. 82-30

Center for Multivariate Analysis
University of Pittsburgh
Ninth Floor, Schenley Hall
Pittsburgh, PA 15260

F49620-82-K-0001

[1]This work is sponsored by the Air Force Office of Scientific Research under Contract [redacted]. Reproduction in whole or in part is permitted for any purpose of the United States Government.

[2]On leave from Fachbereich 17 (Math.), Gesamthochschule, Postfach 101380, D-3500 Kassel, F.R. Germany.

LINEAR SUFFICIENCY AND SOME APPLICATIONS IN MULTILINEAR ESTIMATION

by

Hilmar Drygas

Summary

In the linear model $Y = X\beta + u$ the question arises when a linear transformation $z = Ly$ contains all information of the linear model. This problem was solved by Baksalary and Kala (Annals 1981), Drygas (Sankhyā, forthcoming) and J. Müller (Ph.D. thesis, Kassel 1982). As an application we consider the estimation of the variance of the observations, its skewness and its kurtosis. This is done by considering so-called derived models. (Anscombe, Pukelsheim, Kleffe). Linear sufficient statistics are derived for these problems.

Key words and phrases: Linear models, tensor-products, symmetric tensors, variance, skewness, kurtosis, multilinear estimation, linearly sufficient statistics.

LINEAR SUFFICIENCY AND SOME APPLICATIONS IN MULTILINEAR ESTIMATION

Hilmar Drygas

1. Linearly sufficient statistics in linear models

The concept of linear sufficiency goes back to work by Baksalary and Kala [2], Drygas [6] and J. Müller [11]. Since it is needed in a coordinate- free form we will give it here in this form. As usual a linear model is described by a statistical field (Ω, F, P) and a random H-valued vector y, H an inner-product space, such that

(1.1) $$E_p y \in L \quad \forall P \in P$$

(1.2) $$\mathrm{Cov}_p y \in Ⓗ \quad \forall P \in P$$

This setup is also called the model M(L, Ⓗ). L will in general be a linear manifold and Ⓗ a cone of n.n.d. matrices (or operators). In this paper we will only be concerned with the case Ⓗ $= \{\sigma^2 Q: \sigma^2 \geq 0\}$.

If the model M(L, Ⓗ) is given then a linear inhomogeneous transformation d + Gy, G a linear mapping from H to H, is called BLUE (Best linear unbiased estimator) of Ey if it is unbiased and has smallest covariance-matrix (covariance-operator) in the class of all linear unbiased estimator of Ey. d+ Gy is BLUE of Ey iff

(i) $d = (I-G)l \ \forall\, l \in L$, (ii) $Gf = f \ \forall\, f \in F = L - L$ and

(iii) $GQx = 0 \ \forall\, x \in F^{\perp} \wedge \forall\, Q \in$ Ⓗ.

A BLUE must not exist, but it exists in the case $H = \{\sigma^2 Q;\ \sigma^2 \geq 0\}$ since $F \cap QF^{\perp} = \{0\}$. (See e.g., Drygas [5]).

1.1 __Definition__: Let $A_0 y = c + Ay$. Then $A_0 y$ is called linearly sufficient if there is a BLUE of Ey which is a linear function of $A_0 y$.

1.2 __Theorem__: $A_0 y$ is linearly sufficient if and only if $F \subseteq \mathrm{im}(W A^*)$, where $W = Q + cP_F$ is such that $c \geq 0$ and $F \subseteq \mathrm{im}(W)$. ($P_F$ is the orthogonal projection onto F, A^* is the adjoint mapping of A).

__Proof__: 1. First assume that $F \subseteq \mathrm{in}(WA^*)$. We consider the equation $BAP_F = P_F$. We claim that this equation possesses a solution. This euqation is equivalent to $P_F = P_F A^* B^*$ or $F \subseteq \mathrm{in}(P_F A^*)$ which again is equivalent to $(AP_F)^{-1}(0) \subseteq F^{\perp}$. Therefore let $AP_F x = 0$, then $P_F x = WA^* b$ for some b and $AP_F x = AWA^* b = 0$, implying $WA^* b = P_F x = 0$.

Now let Hz be a BLUE of Ez in the model $M(AF, AQA^*)$. Then for $l \in L$

$$(1.3) \qquad (I - BHA)\, l + BAy$$
$$= (I - BHA)\, l + BGHc + BHA_0 y$$

is BLUE of Ey, if $BAP_F = P_F$. Indeed, if $l \in L$ and $y = Qw$, $w \in F^{\perp}$, then $AQw = AWw = AWA^* v$ for some v. Since $F \subseteq \mathrm{in}(WA^*)$ is equivalent to $(AW)^{-1}(0) \subseteq F^{\perp}$, $AW(w - A^* v) = 0$ implies $w - A^* v \in F^{\perp}$, i.e., $A^* v \in F^{\perp}$ or $v \in A^{*-1}(F^{\perp}) = (AF)^{\perp}$. Thus $HAWA^* v = 0$. Thus

$$(1.4) \qquad (I - BHA)l - Ghc + GHA_0(l + Qw)$$
$$= l\,,$$

proving the BLUE-property.

2. Let G_oA_oy+d be BLUE of Ey in $M(L,\{Q\})$. Then

(1.5) $$G_0Af = 1 \ \forall\ 1 \in F = L-L,\ G_0AQF^{\perp} = 0.$$

We show $(AW)^{-1}(0) \subseteq F^{\perp}$ which is equivalent to $F \subseteq$ in (WA^*). Let $AWt = 0$, where $Wt = Qa + f$, $a \in F^{\perp}$, $f \in F$. Then $G_0AWt = 0 = f$ implying $Wt = Qa = Wa, (t-a) \in W^{-1}(0) \subseteq F^{\perp}$ (since $F \subseteq \operatorname{im} W$). Thus $t = a + (t-a) \in F^{\perp}$.

Q.E.D.

1.3 <u>Definition</u>: Let $z = A_0y$ be linearly sufficient. Then z is called linearly minimal sufficient if for any $z_1 = A_1y$ which is linearly sufficient, there exists a B_1 such that $z = B_1z_1$ almost surely mod P.

1.4 <u>Theorem</u>: $z = A_0y \equiv c + Ay$ is linearly minimal sufficient if and only if

(1.5) $$F = \operatorname{im}(WA^*)$$

The proof goes along the lines of a similar proof in Drygas [6].

2. Computation of expectation and covariance for multilinear expressions

In this paragraph we are assuming that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are independent (at least up to a required order concerning the computation of moments) random variables with expectation zero and existing moments up to some required order. The moments $E(\varepsilon_i^k)$ are assumed to be equal for all i. Thus $\varepsilon_1,\ldots,\varepsilon_n$ behave - at least what the moments up to a certain order is concerned - as independently identically distributed random variables.

Let $\varepsilon = (\varepsilon_1,\ldots,\varepsilon_n)'$ and A be a symmetric n×n-matrix. Then

$$(2.1) \qquad E(\varepsilon' A \varepsilon) = E(\mathrm{tr}\,(A\,\varepsilon\varepsilon')) = E(\mathrm{tr}\,(A\,\sigma^2 I_n)) = \sigma^2\,\mathrm{tr}\,A.$$

The computation of $E(\varepsilon' A \varepsilon)^2$ or $\mathrm{Var}(\varepsilon' A \varepsilon)$ is tedious but it is usually considered as"elementary and straightforward". However, in the last years attempts have been made to make such computations more efficient. We mention in this context mainly the paper by J. Kleffe [10] who has elaborated an approach originally adapted by Balestra [3] and Neudecker [12].

Let $A = (a_{ij},\ i,j = 1,2,\ldots,n)$. Then for computing $E(\varepsilon' A \varepsilon)^2$ evidently

$$(2.2) \qquad E((\sum_{i,j} a_{ij}\varepsilon_i\varepsilon_j)(\sum_{k,l} a_{kl}\varepsilon_k\varepsilon_l)) = \sum_{i,j,k,l} a_{ij}a_{kl}\,E(\varepsilon_i\varepsilon_j\varepsilon_k\varepsilon_l)$$

is needed. Since $(a_{ij}a_{kl}) = A \otimes A$, where $\otimes$ denotes the Kronecker-product $A \otimes B = (a_{ij}B)$, evidently

$$(2.3) \qquad E(\varepsilon' A \varepsilon)^2 = \mathrm{tr}((A \otimes A)\,E(\varepsilon\varepsilon' \otimes \varepsilon\varepsilon')).$$

This formula has been obtained by Balestra [3], Neudecker [12] and Kleffe [10] via a different technique. However, the formula (2.3) does not yet help very much. Let us therefore rewrite (2.2), (2.3) in the form

(2.3a) $$E(\varepsilon' A\varepsilon)^2 = \sum_{i,j} a_{ij}\left(\sum_{k,l} a_{kl}\, E(\varepsilon_k\varepsilon_l\varepsilon_i\varepsilon_j)\right).$$

If we denote the $n\times n$ - matrix $E(\varepsilon_k\varepsilon_l\varepsilon\varepsilon') = E(\varepsilon_k\varepsilon_l\varepsilon_i\varepsilon_j)$ by ψ_{kl}, then evidently

(2.4) $$E(\varepsilon' A\, \varepsilon)^2 = \operatorname{tr}(A \cdot \sum_{k,l} a_{kl}\, \psi_{kl}).$$

This is the formula obtained by Kleffe [10]. If $k = l$, then by independence $\psi_{kl} = \psi_{kk} = E(\varepsilon_i^4)e_{kk} + \sigma^4 \sum_{\delta\neq k} e_{\delta\delta}$ by if we denote the matrix $e_i\, e_j' - e_i$ the i-th unit-vector in $\mathbb{R}^n$ - by e_{ij}. Let $E(\varepsilon_i^4) = \beta\, \sigma^4$. Then

(2.5) $$\psi_{kk} = \sigma^4\{(\beta-1)\, e_{kk} + \sum_{\delta} e_{\delta\delta}\} = \sigma^4\{(\beta-1)e_{kk} + I_n\}$$

is obtained. Similarly we get for $k \neq l$, that by independence

(2.6) $$\psi_{kl} = \sigma^4(e_{kl} + e_{lk}) .$$

Finally by symmetry of A

(2.7) $$\sum_{k,l} a_{kl}\psi_{kl} = \sigma^4\{(\sum_{k=1}^{n} a_{kk})(I_n) + \sum_{k}(\beta-1)a_{kk}e_{kk} + 2\sum_{k\neq l} a_{kl}e_{kl}\}$$

$$= \sigma^4\{(\operatorname{tr} A)\, I_n + (\beta-3)\operatorname{diag} A + 2\,A\} ,$$

where $\operatorname{tr} A = \sum_{k=1}^{n} a_{kk}$ and diag A is the diagonal matrix with the same diagonal-matrix as A. Since $(\operatorname{tr} A)^2 = \operatorname{tr}((\operatorname{tr} A)I_n \cdot A)$, evidently

(2.8) $$(\operatorname{Cov} \varepsilon\varepsilon')\, A = 2\,A + (\beta-3)\operatorname{diag} A,$$

as is well-known (Hsu [8], Drygas [4]).

The method developed by Kleffe can readily be extended to the computation of variances for p-fold Kronecker-products. We consider the $\mathbb{R}^{n^p}$ as the set of collection of real numbers $(a_{i_1,\dots,i_p}, i_1,\dots,i_p = 1,\dots,n)$ which are lexikographically ordered. An element $a = (a_{i_1,\dots i_p})$ will also be called a p-fold tensor. If $b \in \mathbb{R}^n$, then $b \otimes P =: (b^{\otimes p-1}) \otimes b$ is evidently a p-fold tensor with elements $b_{i_1,\dots,i_p} = b_{i_1} b_{i_2} \dots b_{i_p}$. In $\mathbb{R}^{n^p}$ we introduce the usual inner product

$$(2.9) \qquad \langle a,b\rangle = \sum_{i_1,\dots i_p=1}^{n} a_{i_1\dots i_p} b_{i_1\dots i_p} .$$

A tensor $a = (a_{i_1,\dots,i_p})$ is called symmetric, if

$$(2.10) \qquad a_{\pi(i_1)\dots\pi(i_p)} = a_{i_1\dots i_p}$$

for any permutation $\pi \in S_p$. Evidently $b^{\otimes p}$ is a symmetric tensor. The projection on the set of symmetric matrices is given by the symmetrizer π_S:

$$(2.11) \qquad (\pi_S a)_{i_1\dots i_p} = \frac{1}{p!} \sum_{\pi\in S_p} a_{\pi(i_1)\pi(i_2)\dots\pi(i_p)} .$$

We consider $\varepsilon^{\otimes 3}$ and $\varepsilon^{\otimes 4}$. Evidently

$$(2.12) \qquad E\langle a,\varepsilon^{\otimes 3}\rangle = \Big(\sum_{i=1}^{n} a_{iii}\Big) E(\varepsilon_1^3)$$

Similarly, if a is symmetric

$$(2.13) \qquad E(\langle a,\varepsilon^{\otimes 4}\rangle) = \sum_{i,j,k,l} a_{ijkl} E(\varepsilon_i\varepsilon_j\varepsilon_k\varepsilon_l) =$$

$$\sigma^4(\beta-3) \sum_i a_{iiii} + 3\,\sigma^4 \sum_{i,j} a_{iijj} ,$$

where again $E(\varepsilon_i^4) = \beta\sigma^4$. Since $\langle a,\varepsilon^{\otimes 4}\rangle = \langle a,\pi_S\, \varepsilon^{\otimes 4}\rangle = \langle\varepsilon^{\otimes 4},\pi_S a\rangle$ the restriction to symmetric a is not essential. We will come back to this at the end of the paragraph.

What the computation of the covariance-operator of $\varepsilon^{\otimes 3}$ and

$\varepsilon^{\otimes 4}$ is concerned it is hardly possible to get simple expressions without additional assumptions. Therefore we will assume in the sequel that ε is quasi-normally distributed, i.e., that the moments up to order 6 and 8, respectively, coincide with the normal moments. This means that for $p = 3$ we assume that $E(\varepsilon_i^4) = E(\varepsilon_1^4) = 3\sigma^4$, $E(\varepsilon_i^6) = E(\varepsilon_1^6) = 15\sigma^6$ and for $p = 4$ additionally $E(\varepsilon_i^8) = 105\sigma^8$ is assumed to hold. To compute $E(\langle a, \varepsilon^{\otimes p}\rangle^2)$ evidently

$$(2.14)\quad \sum_{i_1,\dots,i_p,j_1,\dots,j_p=1}^{n} a_{i_1\dots i_p} a_{j_1\dots j_p} \varepsilon_{i_1}\varepsilon_{i_2}\dots\varepsilon_{i_p}\varepsilon_{j_1}\dots\varepsilon_{j_2}\dots\varepsilon_{j_p}$$

has to be computed. This may also be written as $\langle a, Va\rangle$, where V is some operator. Evidently (2.14) is equal to

$$(2.15)\quad \sum_{i_1\dots i_p=1}^{n} a_{i_1}\dots a_{i_p}\Big(\sum_{j_1\dots j_p} a_{j_1\dots j_p} E(\varepsilon_{j_1\dots j_p}\varepsilon_{i_1},\dots,\varepsilon_{i_p})\Big)$$

implying that

$$(2.16)\qquad Va = \sum_{j_1\dots j_p=1}^{n} a_{j_1\dots j_p}\,\psi_{j_1\dots j_p},$$

where $\psi_{j_1\dots j_p} = E(\varepsilon_{j_1}\dots\varepsilon_{j_p}\varepsilon^{\otimes p})$.

We will compute Va for $p = 3$ via formula (2.16) and for $p = 4$ by just computing $E(\varepsilon_{j_1}\dots\varepsilon_{j_p}\varepsilon_{i_1}\dots\varepsilon_{i_p})$. This will allow a comparison of the two methods.

To do the computations for $p = 3$ evidently $\psi_{\alpha\beta\gamma}$, $\psi_{\alpha\alpha\beta}$ and $\psi_{\alpha\alpha\alpha}$ $(\alpha \neq \beta \neq \gamma)$ have to be considered. We denote by $e_{\alpha\beta\gamma}$ the tensor having a 1 at place (α, β, γ) and zero elsewhere. Let, moreover, $I_\beta = \sum_{\alpha=1}^{n} a_{\alpha\alpha\beta}$. Then we get for symmetric a:

(2.17) $$\psi_{\alpha\beta\gamma} = \sum_{\pi\in S_3} e_{\pi(\alpha)\pi(\beta)\pi(\gamma)} = 6\,\pi_S e_{\alpha\beta\gamma}$$

(2.18) $$\psi_{\alpha\alpha\beta} = 3\sigma^6 e_{\beta\beta\beta} + \sigma^6 \sum_{\delta\neq\alpha,\beta} (e_{\delta\delta\beta} + e_{\beta\delta\delta} + e_{\delta\beta\delta})$$

$$+ 3\sigma^6 (e_{\alpha\alpha\beta} + e_{\alpha\beta\alpha} + e_{\beta\alpha\alpha}) = 3\sigma^6 \pi_S I_\beta + \, , \sigma^6 \pi_S e_{\alpha\alpha\beta}.$$

(2.19) $$\psi_{\alpha\alpha\alpha} = 15\sigma^6 e_{\alpha\alpha\alpha} + 3\sigma^6 \sum_{\beta\neq\alpha} (e_{\beta\beta\alpha} + e_{\alpha\beta\beta} + e_{\beta\alpha\beta})$$

$$= 6\sigma^6 e_{\alpha\alpha\alpha} + 9\sigma^6 \pi_S I_\alpha.$$

Finally

(2.20) $$Va = \sum_{\alpha,\beta,\gamma} a_{\alpha\beta\gamma}\psi_{\alpha\beta\gamma} = \sigma^6\{6 \sum_{\alpha\neq\beta\neq\delta} a_{\alpha\beta\delta} e_{\alpha\beta\delta}$$

$$+ 9 \sum_\beta \Big(\sum_{\alpha\neq\beta} a_{\alpha\alpha\beta}\Big) \pi_S I_\beta + 6\sigma^6 \sum_{\alpha\neq\beta} a_{\alpha\alpha\beta}(e_{\alpha\alpha\beta} + e_{\alpha\beta\alpha} + e_{\beta\alpha\alpha})$$

$$+ 6 \sum_\alpha a_{\alpha\alpha\alpha} e_{\alpha\alpha\alpha} + 9 \sum_\alpha (\pi_S I_\alpha) a_{\alpha\alpha\alpha}\}$$

$$= \sigma^6\{6 \sum_{\alpha,\beta,\gamma} a_{\alpha\beta\gamma} e_{\alpha\beta\gamma} + 9 \sum_{\alpha,\beta} a_{\alpha\alpha\beta}\pi_S I_\beta\}$$

$$= \sigma^6\{6a + 9 \sum_{\beta=1}^{n} (\mathrm{tr}_\beta\, a)\pi_S I_\beta\},$$

where $\mathrm{tr}_\beta\, a = \sum_{\alpha=1}^{n} a_{\alpha\alpha\beta}$. Evidently also

(2.21) $$Va = \sigma^6\{6a + 9 \sum_{\beta=1}^{n} (\pi_S I_\beta \circ \pi_S I_\beta)\},$$

where $a \circ b$ denotes the outer product defined by $(a \circ b)c = \langle b,c\rangle a$.

This follows from $\langle a, \pi_S I_\beta \rangle = \mathrm{tr}_\beta a$, if a is symmetric. It is also true, that

$$(2.22) \quad \sum_{\beta=1}^{n} (\pi_S I_\beta \circ \pi_S I_\beta) = \pi_S (I_n \otimes \mathrm{vec}(I_n)(\mathrm{vec}\, I_n)')\pi_S ,$$

the representation found by Pukelsheim in [14]. This representation will not be used in this paper.

We now assert, that for symmetric $a = (a_{ijkl})$

$$(2.23) \quad E(\langle a, \varepsilon^{\otimes 4} \rangle^2) = \sum_{i,j,k,l,r,s,t,u=1}^{n} a_{ijkl} a_{rstu} E(\varepsilon_i \varepsilon_j \varepsilon_k \varepsilon_l \varepsilon_r \varepsilon_s \varepsilon_t \varepsilon_u)$$

$$= \sigma^8 \{ 24 \sum_{i,j,k,l=1}^{n} a_{ijkl}^2 + 72 \sum_{i,j,k,l=1}^{n} a_{iijk} a_{jkll}$$

$$+ 9 \sum_{i,j,k,l=1}^{n} a_{iijj} a_{kk} a_{ll} \} .$$

Indeed, under quasi-normality, $E(\varepsilon_i \varepsilon_j \varepsilon_k \varepsilon_l \varepsilon_r \varepsilon_s \varepsilon_t \varepsilon_u)$ vanishes if some ε_α appears an uneven number of times. Therefore only the cases ε_i^8, $\varepsilon_i^6\ \varepsilon_j^2$, $\varepsilon_i^4\ \varepsilon_j^4$, $\varepsilon_i^4\ \varepsilon_j^2\ \varepsilon_k^2$ and $\varepsilon_i^2 \varepsilon_j^2 \varepsilon_k^2 \varepsilon_l^2$ are to be considered. If all indices i,j,k,l are different from each other then surely the sums reported in (2.23) will appear. The factor 24, 72 and 9 arise from careful combinatorial considerations and the fact that a is symmetric. Note that some combinations are covered by the summation. If $i = j = k = l$, then the subsum in (2.22) is equal to

$$(2.24) \qquad 105\, \sigma^8 \sum_{i=1}^{n}{}' a_{iiii}^2 = \sum_{i=1}^{n} a_{iiii}^2\, E(\varepsilon_i^8).$$

Now consider the 6 cases $i = j$, $i = k$, $i = l$, $j = k$, $j = l$ and $k = l$. Then by symmetry the corresponding subsum in (2.23) is equal to

(2.25) $\sigma^8\{18\ \Sigma a_{iiii}a_{jjkk} + 288\ \Sigma a_{iiij}a_{ijkk}$

$$+ 216\ \Sigma a_{iijk}^2 + 108\ \Sigma\ a_{iijj}a_{iikk}\}$$

$$= 3\ \sigma^8\{6\ \Sigma a_{iiii}a_{jjkk} + 96\ \Sigma a_{iiij}a_{ijkk}$$

$$+ 72\ \Sigma a_{iijk}a_{iijk} + 36\ \Sigma a_{iijj}a_{iikk}\}$$

In view of $E(\varepsilon_i^4) = 3\ \sigma^4$ and symmetry this is just the set of all possible summands occurring with factor $E(\varepsilon_i^4\varepsilon_j^2\varepsilon_k^2)$. Again careful considerations are necessary to establish the combinatorial numbers 6,96,72 and 36. Finally we get for the seven cases $i = j = k$; $i = j = l$; $i = k = l$; $j = k = l$; $i = j$, $k = l$; $i = k$, $j = l$ and $i = j$, $k = l$ as subsum of (2.23)

(2.26) $6^8\{180\ \Sigma a_{iiii}a_{iijj} + 240\ \Sigma a_{iiij}^2$

$$+ 9\ \Sigma a_{iiii}a_{jjjj} + 144\ \Sigma a_{iiij}a_{ijjj} + 162\ \Sigma a_{iijj}^2\}$$

$$= 15\sigma^8\{12\ \Sigma a_{iiii}a_{iijj} + 16\ \Sigma a_{iiij}^2\}$$

$$+ 9\sigma^8\{\Sigma a_{iiii}a_{jjjj} + 16\ \Sigma a_{iiij}a_{ijjj} + 18\ \Sigma a_{iijj}^2\}.$$

In view of $E(\varepsilon_i^6) = 15\sigma^6$, $E(\varepsilon_i^4) = 3\sigma^4$ the first sum belongs due to symmetry of a to all terms where $E(\varepsilon_i^6\varepsilon_j^2)$ appear. The second sum belongs to all terms where $E(\varepsilon_i^4\varepsilon_j^4)$ occurs. Again, careful reasoning is necessary to determine the combinatorial factor 12, 16, 1, 16 and 18.

Since the last term in (2.22) is evidently equal to $9\sigma^8(\Sigma a_{iijj})^2 = [E(\langle a,\varepsilon^{\otimes 4}\rangle)]^2$ it follows that for symmetric a:

(2.27) $\quad \mathrm{Var}(\langle a,\varepsilon^{\otimes 4}\rangle) = \sigma^8\{24\ \Sigma a_{ijkl}^2 + 72\ \Sigma a_{iijk}a_{jkll}\}$

Define the tensor $e_{\alpha\beta\gamma\delta}$, which has unity at place $(\alpha,\beta,\gamma,\delta)$ and zero elsewhere. Let $I_{jk} = \sum_{\alpha=1}^{n} e_{jk\alpha\alpha}$ and $\mathrm{tr}_{j,k}a = \sum_{\alpha=1}^{n} a_{jk\alpha\alpha}$. Then evidently

(2.28) $\quad \mathrm{Var}(\langle a,\varepsilon^{\otimes 4}\rangle) = \sigma^8\{\langle a, 24a + 72\sum_{j,k=1}^{n}(\mathrm{tr}_{jk}a)\pi_S I_{jk}\rangle\}.$

This shows that for symmetric a evidently

(2.29) $\quad \mathrm{Cov}(\varepsilon^{\otimes 4}) = \sigma^8\{24\ I + 72\sum_{j,k=1}^{n}(\pi_S I_{jk} \circ \pi_S I_{jk})\}$

where $(a \circ b)$ again denotes the outer product: $(a \circ b)c = \langle b,c\rangle a$. (2.28) can also be written as

(2.30) $\quad \sigma^8\{24\,I + 72\{\pi_S(I_n \otimes I_n \otimes (\mathrm{vec}\,I_n)(\mathrm{vec}\,I_n)' \otimes (\mathrm{vec}\,I_n)(\mathrm{vec}\,I_n)'\pi_S\}$

the representation given in Pukelsheim [14]. This representation will not be used here.

A final remark of this paragraph concerns the covariance-opeartor (2.28). This formula is only correct if it is really considered as a covariance-operator, restricted to the symmetric tensors. It is not identical with the covariance-matrix. Let us assume we have computed the covariance-matrix $E(\varepsilon^{\otimes 2}(\varepsilon^{\otimes 2})') = C$. From (2.4)-(2.6) we get in the quasi-normal case $Ca = \sigma^4\{(a_{ij}) + (a_{ji}) + (\mathrm{tr}\ a)I_n\}$. Since in general $\mathrm{vec}(bb') = b \otimes b$ it follows that $E(\varepsilon^{\otimes 4}) = \mathrm{vec}(E(\varepsilon^{\otimes 2}(\varepsilon^{\otimes 2})') = \mathrm{vec}(C)$. Denote by e_{ij} the vector in $\mathbb{R}^{n^2}$ having 1 at place $i\,n + j$ and zero elsewhere, then it is easily checked that vec(C) is different from $3\sigma^4$ $(e_{11}\ldots e_{1n}\ldots e_{n1}\ldots e_{nn})$, but equality is obtained when the two matrices are applied to symmetric tensors. In so far the assertion in Pukelsheim [14] claiming that (2.29) is the covariance-matrix is wrong.

3. Linear sufficient statistics in multilinear estimation

We consider the linear model $Ey \in L \subseteq \mathbb{R}^n$, $\text{Cov}\, y = \sigma^2 I_n$ as described in section 1 of this paper. Let $F = L-L$ and $P_F y$ denote the orthogonal projection of y onto F. Then $G_o y = l + P_F(y-l)$, $l \in L$ is the unique BLUE of Ey. We consider

$$(3.1) \qquad u = \sigma^{-1}(y-Ey), \; z = (I-P_F)(y-l) = \sigma(I-P_F)u.$$

The quantity

$$(3.2) \quad V = zz' = (I-P_F)(y-l)(y-l)'(I-P_F) = \sigma^2(I-P_F)uu'(I-P_F)$$

is a random element with values in the set H of all symmetric n×n-matrices A satisfying $Af = 0 \; \forall \; f \in F$. Let $M = (I-P_F)$, then $A \in H$ iff $MAM = A$ (Drygas [4]). In H the inner product $\langle A,B\rangle = tr(AB)$ is used. Since $\langle zz',A\rangle = \sigma^2(u'Au)$ and by (2.1), (2.8)

$$(3.3) \qquad E(u'Au) = tr(A) = tr(MA)$$

$$(3.4) \qquad \text{Var}(u'Au) = tr([2A + (\beta-3)\,\text{diag}\,A]\cdot A) = tr([2A + (\beta-3)\,M\cdot\text{diag}\,A\cdot M]A),$$

we get that $V = zz'$ follows the linear model

$$\text{(3.5)} \qquad E\,V = \sigma^2 M, \quad \text{Cov}\,V = \sigma^4\{2A + (\beta-3)M \text{ diag } A\,M\},$$

if considered as H-valued random element.

Besides the mapping $\text{Diag } A = (a_{ij}\delta_{ij})$ which is evidently self-adjoint we consider the linear mapping diag: $\mathbb{R}^n \to \mathbb{R}^{n\times n}$ defined by $\text{diag }(x_1,\ldots,x_u)' = (\delta_{ij}x_i)$. Evidently the adjoint is $\text{diag}^*(a_{ij}) = (a_{11},\ldots,a_{nn})'$. If A and B are two n×n-matrices then the Hadamard-product $A * B$ is defined by $(A * B) = (a_{ij}b_{ij})$.

3.1 <u>Theorem</u>: Let $M \neq 0$. Then $\text{tr}(MV) = \text{tr}(V)$ is a linearly sufficient statistic in the model (3.5) iff the Hsu-condition $\sigma^4(\beta-3)(M * M)\, m = \rho\, m$, where $\rho = (\beta-3)\sigma^4 \text{ tr}(M * M)/(\text{tr } M)$ is met. In this case $\text{tr}(MV) = \text{tr } V$ is also linearly minimal sufficient.

<u>Proof</u>: Consider the linear mapping $AV = \text{tr}(MV)$. This is a mapping from H to $\mathbb{R}$. The adjoint mapping of this mapping is $A^*\lambda = \lambda M$, $\lambda \in \mathbb{R}$. A is linearly sufficient iff

$$\text{(3.6)} \qquad \{\lambda M\} \subseteq \{\lambda\, W M\} = \text{im}(WA^*),$$

where

$$\text{(3.7)} \qquad W = \text{Cov } V + (\text{tr } M)^{-1}\,(M \circ M),$$

since $(\text{tr } M)^{-1} M \circ M$ is the orthogonal projection onto $\{\lambda M\}$. This is the case iff

$$\text{(3.8)} \qquad WM = 2\sigma^4 M + \sigma^4(\beta-3)\, M(\text{Diag } M)\, M + M = \rho M$$

for some $\rho \neq 0$. This means that $\sigma^4(\beta-3)\, M(\text{Diag } M)\, M = \alpha M$ for some $\alpha \in \mathbb{R}$. By taking traces on both sides of the last equation

$\alpha = (\mathrm{tr}\ M)^{-1}\ \alpha^4(\beta-3)\ \mathrm{tr}(M\ \mathrm{diag}\ M) = (\mathrm{tr}\ M)^{-1}\ \sigma^4(\beta-3)\ \mathrm{tr}(M * M)$. Since $\mathrm{tr}(M * M) \leq \mathrm{tr}(M)$ and $\beta \geq 1$ it follows that $\rho = 2\sigma^4 + \alpha + 1 \geq 1 > 0$. But by Hsu's theorem (Drygas-Hupet [7], Pukelsheim [13], Khatri [9]), $M\ \mathrm{diag}\ x\ M = 0$ is equivalent to $(M * M)x = 0$. This finishes the proof of the theorem since the assertion concerning linear minimal sufficiency is now obvious.

3.2 <u>Theorem</u>. Let $M \neq 0$. Then $\mathrm{diag}^*\ V = (v_{11},\dots,v_{nn})'$ is a linearly sufficient statistic.

<u>Proof</u>: 1) Since we consider V as an element of H, diag^* is to be considered as mapping from H to $\mathbb{R}^n$. The adjoint $((\mathrm{diag})^*)^*$ of this mapping is not diag but M diag M, since for $A \in H$

$$(3.9)\qquad \mathrm{tr}\,(M\ \mathrm{diag}\, x M \cdot A) = \mathrm{tr}(\mathrm{diag}\, x \cdot A) = x'(\mathrm{diag})^* A$$

and $M\ \mathrm{diag}\, x\ M \in H$.

2) Two cases have to be distinguished. Either there is an element $A \in H$ such that $(\mathrm{Cov}\ V)A = M$ or there is an element $A \in H$ such that $\mathrm{tr}(AM) \neq 0$ and $(\mathrm{Cov}\ V)A = 0$. (The latter case can only occur if $\beta = 1$.) This follows from $\mathrm{im}(Q) = (Q^{-1}(0))^{\perp}$, if Q is self-adjoint.

In the first case theorem 1.2 tells us that we can choose $W = \mathrm{Cov}\ V$, while in the second case W will be chosen equal to $\mathrm{Cov}\ V + (\mathrm{tr}\ M)^{-1}(M \circ M)$. In both cases, however, $M \in \mathrm{im}(W(\mathrm{diag}^*))^*)$ has to be proved. Let $M = (\mathrm{Cov}\ V)A = \sigma^4\ \{2A + (\beta-3)\ M\ \mathrm{Diag}\ A M\}$, $A \in H$. This implies at first that σ^4 can't vanish. Therefore

(3.10) $A = (2\sigma^4)^{-1} M(\text{diag } I_n - \sigma^4(\beta-3)\text{diag Diag}^* A) M = M \text{ diag } x \, M$

with $x = (2\sigma^4)^{-1}(1_n - \sigma^4(\beta-3) \text{ Diag}^* A)$, $1_n = (1_1, \ldots, 1)'$. Thus $M = (\text{Cov } V) (\text{diag}^*)^* x$ and diag^* is linearly sufficient.

In the second case $(\text{Cov } V) A = 0$, $WA = (\text{tr } M)^{-1} \text{ tr}(MA)M \neq 0$, implying $WB = M$, $B = (\text{tr}(MA))^{-1}(\text{tr } M)M$ and

(3.11) $$(\text{Cov } V)B = 2\sigma^4 B + \sigma^4(\beta-3) M \text{ diag Diag}^* BM = 0.$$

Thus if $\sigma \neq 0$, $B = -2^{-1}(\beta-3) M \text{ diag Diag}^* BM = M \text{ diag } x M$, $x = -2^{-1}(\beta-3) \text{ Diag}^* B$. If $\sigma = 0$, then evidently $W(\text{diag}^*)^* 1_n = W M \text{ diag } 1_n M = M$. Therefore linear sufficiency is proved in all possible cases, Q.E.D.

We will now consider

(3.11) $$V_i = z^{\otimes i} = M^{\otimes i} z^{\otimes i}, \quad i = 3,4.$$

Since $M^{\otimes i} z^{\otimes i} = (Mz)^{\otimes i} = \sigma^i M^{\otimes i} u^{\otimes i}$ we can apply the results of paragraph 2 for obtaining expectation and covariance-operator of V_i. First of all, note that V_i is a symmetric tensor obeying the equation $M^{\otimes i} V_i = V_i$. Therefore our reference vector-space H will be the set of all symmetric tensors a meeting the equation $M^{\otimes i} a = a$.

We introduce the following notation: Let $a \in \mathbb{R}^n$, $a = (a_1, \ldots, a_n)'$. Then we define $\text{diag}_1 \, a = (a_{i_1} \delta_{i_1 i_2} \cdots \delta_{i_1 i_p}, \; i_1, \ldots, i_p = 1,2,\ldots,n) \in \mathbb{R}^{n^p}$. In general if $a = (a_{i_1 \ldots i_k}) \in \mathbb{R}^{n^k}$ and $p > k$ we define

(3.12) $$\text{diag}_k(a) = (a_{i_1 \ldots i_k} \delta_{i_k i_{k+1}} \cdots \delta_{i_k i_p}) \in \mathbb{R}^{n^p}.$$

With these definitions we evidently get from (2.11) and (2.12):

(3.13) $$E(z^{\otimes 3}) = \sigma^3 E(u_1^3) M^{\otimes 3}\ \mathrm{diag}_1\ 1_n,\ 1_n = (1,\ldots,1)'.$$

(3.14) $$E(z^{\otimes 4}) = \sigma^4 M^{\otimes 4}\ \{(\beta-3)\ \mathrm{diag}_1\ 1_n + 3 \sum_{i=1}^{n} \pi_S\ I_{ii}\}.$$

The covariance-operator, defined as mappings from H to H, are found in the quasi-normal case (use $M^{\otimes i} a = a!$) to be equal to

(3.15) $$\mathrm{Cov}(z^{\otimes 3}) = \sigma^6\{6I + 9\ M^{\otimes 3} \sum_{\beta=1}^{n} (\pi_S I_\beta \circ \pi_S I_\beta)\},$$

(3.16) $$\mathrm{Cov}(z^{\otimes 4}) = \sigma^8\{24I + 72\ M^{\otimes 4} \sum_{j,k=1}^{n} \pi_S I_{jk} \circ \pi_S I_{jk}\}.$$

This model has intensively been studied in Pukelsheim [14]. Since the covariance-operator is only computed under quasi-normality the estimators derived from linear model theory are only locally best (linear) unbiased estimators. Pukelsheim's investigation was suggested by a paper by Anscombe [1], who used $\mathrm{diag}_1^*\ z^{\otimes 3} = z^{*3}$ and $\mathrm{diag}_1^*\ z^{\otimes 4} = z^{*4}$ to obtain estimators of $E(u_1^3)$ and $E(u_1^4)$, respectively. Pukelsheim showed that these estimators are not even locally best. Using $z^{\otimes 3}$ and $z^{\otimes 4}$, respectively, means the consideration of linear combinations of $z_i z_j z_k$ $(i,j,k = 1,\ldots,n)$ and of $z_i z_j z_k z_l$ $(i,j,k,l = 1,\ldots,n)$, respectively. But we will show now that it is enough to consider only $z_i^2 z_j$ $(i,j = 1,\ldots,n)$ and $z_i^2 z_j z_k$ $(i,j,k = 1,\ldots,n)$, respectively. Evidently $(z_i^2 z_j) = \mathrm{diag}_2(z_i z_j z_k) = z^{*2} \otimes z$. Similarly, $(z_i^2 z_j z_k) = z^{*2} \otimes z^{\otimes 2} = \mathrm{diag}_3\ z^{\otimes 4}$. We will prove that these statistics are linearly sufficient.

3.3 Theorem. a) $z \otimes z^{*2} = \mathrm{diag}_2^*\ z^{\otimes 3}$ is linearly sufficient in the model described by (3.13), (3.15).

b) $z^{\otimes 2} \otimes z^{*2} = \mathrm{diag}_3^* z$ is linearly sufficient in the model described by (3.14), (3.16).

<u>proof</u>: a) Since our reference vector-space H is the set of all symmetric tensor a from $\mathbb{R}^{n^3}$ meeting $M^{\otimes 3} a = a$, $(\mathrm{diag}_2^*)^*$ has to be mapping from $\mathbb{R}^{n^2}$ to H. This mapping is

$$(3.17) \qquad (\mathrm{diag}_2^*)^* = M^{\otimes 3} \pi_S \mathrm{diag}_2$$

This follows since $M^{\otimes 3}\pi_S \mathrm{diag}_2\, b \in H$ and $\langle M^{\otimes 3}\pi_S\, \mathrm{diag}_2\, b, c\rangle = \langle \mathrm{diag}_2\, b, c\rangle = \langle b, \mathrm{diag}_2^*\, c\rangle$ for all $b \in \mathbb{R}^{n^2}$, $c \in H$.

We firstly deal with the case $\sigma^6 = 0$. Let $M \neq 0$, otherwise there is no assertion. Let $W = (M^{\otimes 3}\mathrm{diag}_1\, 1_n \circ M^{\otimes 3}\, \mathrm{diag}_1\, 1_n)(\sum_{i=1}^{n} m_{ii}^3)^{-1}$. W is the orthogonal projection onto $F = \{\gamma\, M^{\otimes 3}\, \mathrm{diag}_1\, 1_n\}$. Since Cov $V_3 = 0$ and $W\, M^{\otimes 3}\, \mathrm{diag}_1\, 1_n = M^{\otimes 3}\, \mathrm{diag}_1\, 1_n = M^{\otimes 3}\, \pi_S\, \mathrm{diag}_2\, \mathrm{vec}(I_n) = (\mathrm{diag}^*)^*\, \mathrm{vec}(I_n)$, (diag_2^*) is clearly linearly sufficient.

Now let $\sigma^6 \neq 0$, then Cov $V_3 = \sigma^6\{6I + \Sigma M^{\otimes 3}\, \Sigma \ldots\}$ is regular or H, since $\langle(\mathrm{Cov}\, V_3)a, a\rangle = 6\, \sigma^6 \langle a, a\rangle + 9\, \sigma^6 \sum_j (\sum_i a_{iij})^2$ vanishes iff $a = 0$. For this reason there is a tensor $a \in H$ such that

$$(3.18) \qquad M^{\otimes 3}\mathrm{diag}_1\, 1_n = 6\, \sigma^6 + 9\, \sigma^6\, M^{\otimes 3} \sum_{j=1}^{n} (\pi_S I_j \circ \pi_S I_j) a$$

Since W can be chosen equal to Cov V_3, our assertion would be proved if we could show that a has the form $M^{\otimes 3}\pi_S \mathrm{diag}_2 b$ for some $b \in \mathbb{R}^{n^2}$. But (3.18) implies that

$$(3.19) \qquad a = (6\sigma^6)^{-1} M^{\otimes 3}\{\mathrm{diag}_1\, 1_n - 9\, \sigma^6 \sum_{j=1}^{n} \langle I_j, a\rangle \pi_S I_j\}$$
$$= M^{\otimes 3}\pi_S \mathrm{diag}_2\, b;$$

where $b = (b_{ij})$ and

(3.20) $$b_{ij} = (6\sigma^6)^{-1}(\delta_{ij} - 9\,\sigma^6 \sum_{i,j=1}^{n} a_{iij}).$$

b) Again, $(dg_3^*)^* = M^{\otimes 4}\pi_S\, dg_3$ can easily be established. Since $\sigma^{-8} \langle(\mathrm{Cov}\ V_4)a,a\rangle = 24\langle a,a\rangle + 72 \sum_{j,k,l} (\sum_l a_{ikll})^2$, $a \in H$ is positive whenever $a \neq 0$, Cov V_4 is regular and $W = \mathrm{Cov}\ V_4$ is a possible choice, if $\sigma^8 \neq 0$ ($\sigma^8 = 0$ can analogously be dealt with as above). Therefore there is an element $a \in H$ such that

(3.21) $$Wa = M^{\otimes 4}\mathrm{diag}_1\, 1_n = \sigma^8\{24a + 72\, M^{\otimes 4} \sum_{j,k} \langle I_{jk},a\rangle \pi_S I_{jk}\}$$

and an element $b \in H$ such that

(3.22) $$Wb = M^{\otimes 4} \sum_i \pi_S I_{ii} = \sigma^8\{24b + 72\, M^{\otimes 4} \sum_{j,k} \langle I_{jk},b\rangle \pi_S I_{jk}\}$$

These two equations can be rewritten as

(3.23) $$\begin{aligned} a &= (24\sigma^8)^{-1}\, M^{\otimes 4}\{\mathrm{diag}_1\, 1_n - 72\,\sigma^8 \sum_{j,k} \langle I_{jk},a\rangle \pi_S I_{jk}\} \\ &= M^{\otimes 4}\{\pi_S\, \mathrm{diag}_3 c_1\}, \end{aligned}$$

(3.24) $$\begin{aligned} b &= (24\sigma^8)^{-1}\, M^{\otimes 4}\{\sum_i \pi_S I_{ii} - 72\,\sigma^8 \sum_{j,k} \langle I_{jk},b\rangle \pi_S I_{jk}\}, \\ &= M^{\otimes 4}\, \pi_S\, \mathrm{diag}_3\, c_2, \end{aligned}$$

where

(3.25) $$c_1 = (\delta_{ij}\delta_{ik} - 72\sigma^8 \sum_{\alpha,\beta} (\sum_\gamma a_{\alpha\beta\gamma\gamma}))/(24\sigma^8)$$

(3.26) $$c_2 = (\delta_{ij} - 72\sigma^8 \sum_{\alpha,\beta,\gamma} a_{\alpha\beta\gamma\gamma})/(24\sigma^8)$$

This shows that $\mathrm{im}(W(\mathrm{diag}_3^*)^*)$ contains the set of possible expectation-values, Q.E.D.

References

[1] Anscombe, F.J. "Examination of residuals". Fourth Berkeley Symp. Math. Statist. Prob. 1, (1961), 1-36.

[2] Baksalary, J. and Kala, R. "Linear transformations preserving best linest unbiased estimators in a general Gauss-Markoff model". Ann. of Statist., (1981), 913-916.

[3] Balestra, P. "Best quadratic unbiased estimators of the variance-covariance matrix in normal regression". J. Econometrics 1, (1973), 17-28.

[4] Drygas, H. "The estimation of residual variance in regression analysis". Math. Operationsf und Statist., 3 (1972), 373-388.

[5] Drygas, H. "Gauss-Markoff estimation for multivariate linear models with missing observations". Ann. Statist. 4 (1976), 779-787.

[6] Drygas, H. "Sufficiency and completeness in the general Gauss-Markoff model". To be published in Sankhya.

[7] Drygas, H. and Hupet, G. "A new proof of Hsu's theorem in regression analysis - a coordinate-free approach". Math. Operationsf. and Stat., Ser. Statist., 8 (1977), 333-335.

[8] Hsu, P.L. "On the best unbiased quadratic estimate of variance". Statist. Res. Memoir 2 (1938), 91-104.

[9] Khatri, C.G. "Minimum Variance Quadratic Unbiased Estimate of Variance under ANOVA Model". Gujarat Statist. Review, 5, (1978), 33-41.

[10] Kleffe, J. "Simultaneous estimation of Expectation and Covariance Matrix in Linear Models". Math. Operations.f.u. Statist., Statistics, 9 (1978), 443-478.

[11] Müller, J. "Suffizienz und Vollständigkeit in linearen Modellen". Ph.D. thesis, Univeristy of Kassel/Germany (1982).

[12] Neudecker, H. "The Kronecker matrix product and some of its applications in econometrics". Statistica Neerlaudica 22 (1968), 69-82.

[13] Pukelsheim, F. "On Hsu's model in regression analysis". Math. Operationsf. Statist., Ser. Statistics, 8 (1977), 323-331.

[14] Pukelsheim, F. "Multilinear estimation of skewness and kurtosis". Metrika 27 (1980), 103-113.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 83-0007	2. GOVT ACCESSION NO.: AD-A125152
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Linear Sufficiency and Some Applications in Multilinear Estimation	5. TYPE OF REPORT & PERIOD COVERED: Technical - October 1982
	6. PERFORMING ORG. REPORT NUMBER: 82-30
7. AUTHOR(s): Hilmar Drygas	8. CONTRACT OR GRANT NUMBER(s): F49620-82-K-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Multivariate Analysis, University of Pittsburgh, Pittsburgh, PA 15260	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research, Department of the Air Force, Bolling Air Force Base, DC 20332	12. REPORT DATE: October 1982
	13. NUMBER OF PAGES: 19
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Linear models, tensor-products, symmetric tensors, variance, skewness, kurtosis, multilinear estimation, linearly sufficient statistics.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In the linear model $Y = X\beta + u$ the question arises when a linear transformation $z = Ly$ contains all information of the linear model. This problem was solved by Baksalary and Kala (Annals 1981), Drygas (Sankhyā, forthcoming) and J. Müller (Ph.D. thesis, Kassel 1982). As an application we consider the estimation of the variance of the observations, its skewness and its kurtosis. This is done by considering so-called derived models. (Anscombe, Pukelsheim, Kleffe.) Linear sufficient statistics are derived for these problems.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)