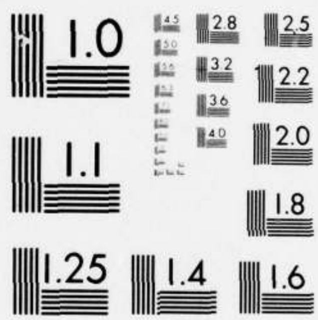
AD-A053 856 TEXAS A AND M UNIV COLLEGE STATION F/G 12/1
A SIMPLE 'SYNTHESIS'-BASED METHOD OF VARIANCE COMPONENT ESTIMAT--ETC(U)
MAR 78 H O HARTLEY, J N RAO, L LAMOTTE DAAG29-77-G-0086
UNCLASSIFIED TR-3 ARO-14209.3-M NL

1 OF 1
AD
A053856

END
DATE
FILMED
6-78
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

ARO 14209.3-M

(12)

AD A053856

ARO-D PROJECT DAAG29-77-G-0086

Technical Report No. 3

A SIMPLE 'SYNTHESIS'-BASED METHOD OF VARIANCE COMPONENT ESTIMATION

by

H. O. Hartley, J. N. K. Rao, and Lynn LaMotte

March 1978

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

BEST AVAILABLE COPY

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: 14209.3-M — ARO — 2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): A Simple 'Synthesis'-Based Method of Variance Component Estimation	5. TYPE OF REPORT & PERIOD COVERED: Technical Report
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): H. O. Hartley, J. N. K. Rao, Lynn LaMotte	8. CONTRACT OR GRANT NUMBER(s): DAAG29-77-G-0086
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Texas A&M University, College Station, Texas 77843	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, P. O. Box 12211, Research Triangle Park, NC 27709	12. REPORT DATE: Mar[redacted]78
	13. NUMBER OF PAGES: 19 22 p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): TR-3	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

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

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

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

In this paper we do not attempt an evaluation of the ever growing methodology in the estimation of variance components. Optimality properties are sometimes achieved at considerable computational efforts. Certain methods depend on a subjective ordering of the components, and if the ordering is unfortunate the method may fail to yield estimates for certain components while with a different ordering all components may well be estimable. The work involved in attempting all possible orderings of the variance components is usually prohibitive. The present method → next page

347 350

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

20. ABSTRACT CONTINUED

cont. achieves optimality properties and is nevertheless computationally simple. In fact it possesses Minque optimality for a particular choice of norm, but also various other optimality properties and necessary and sufficient conditions for estimability associated with Minque simplify considerably. Moreover we are able to derive sufficient conditions for consistency which also provide estimability conditions of a simpler structure. The consistency of our estimators makes them convenient as starting points for a single ML cycle to obtain asymptotically fully efficient estimates.

A SIMPLE 'SYNTHESIS'-BASED METHOD OF VARIANCE COMPONENT ESTIMATION

by

H. O. Hartley*, J. N. K. Rao+ and Lynn LaMotte#

1. Introduction

In this paper we do not attempt an evaluation of the ever growing methodology in the estimation of variance components. (For an excellent summary of the literature up to 1971 see Searle (1971).) Optimality properties are sometimes achieved at considerable computational efforts. A case in point is the M. L. estimation (see Hartley & Rao (1967)) which is still fairly laborious for large data banks in spite of the improvements through the W-transformation (Hemmerle & Hartley (1973)). Similar observations apply to the general case of Minque (C. R. Rao (1971)) recently simplified by Liu & Senturia (1976). Other methods, such as the Henderson 3 Method (Henderson (1953)) or the Abbreviated Doolittle and square root method (see e.g. Gaylor, Lucas and Anderson (1970)) depend on a subjective ordering of the components (such as with the Forward Doolittle procedure) and if the ordering is unfortunate the method may fail to yield estimates for certain components while with a different ordering (not attempted) all components may well be estimable. The work involved in attempting all possible orderings of the variance components is usually prohibitive. The present method achieves optimality properties and is nevertheless computationally simple. In fact it possesses Minque optimality for a particular choice of norm, but also various other optimality properties and necessary and sufficient conditions for estimability associated with Minque simplify considerably (see Section 6). Moreover we are able to derive sufficient conditions for consistency which also provide estimability conditions of a simpler structure (see Appendix). The consistency of our estimators makes them convenient as starting points for a single ML cycle to obtain asymptotically fully efficient estimates.

2. The Mixed ANOVA Model

Employing the currently used notation we write the mixed ANOVA model in the form

*H. O. Hartley, Institute of Statistics, Texas A&M University
+J. N. K. Rao, Carleton University, Ottawa
Lynn LaMotte, Quantitative Management Science, University of Houston

$$y = X\alpha + \sum_{i=1}^{c+1} U_i b_i \qquad (1)$$

where

- y is an $n \times 1$ vector of observations,
- X is an $n \times k$ matrix of known coefficients,
- α is a $k \times 1$ vector of unknown constants,
- U_i is an $n \times m_i$ matrix of 0, 1 coefficients,
- b_i is an $m_i \times 1$ vector of normal variables from $N(0, \sigma_i^2)$.

Specifically $U_{c+1} = I_n$ and b_{c+1} is an n-vector of "error variables". Moreover the design matrices U_i have precisely one value of 1 in each of their rows and all other coefficients 0. We denote by $m = \sum_{i=1}^{c} m_i$ the total number of random levels.

We may assume without loss of generality that

$$X'X = I \qquad (2)$$

for if (2) is not satisfied we may orthogonalize X by a Gram Schmidt orthogonalization process with a consequential reparameterization of α omitting any linearly dependent columns in the Gram Schmidt process. Usually the first column of X is the column vector with all elements = $1/\sqrt{n}$. It is the objective of the method to compute estimates of the variance components σ_i^2 and the vector α.

3. The Present Method

The essence of the present method is to

(a) Select c+1 quadratic forms $Q_j(y)$ in the elements of y.

(b) Use the method of synthesis (Hartley (1967), Rao (1968)) to obtain the coefficients k_{ji} in the formulas for $E(Q_j)$ in the form

$$E(Q_j) = \sum_{i=1}^{c+1} k_{ji}\sigma_i^2 . \tag{3}$$

(c) Estimate σ_i^2 by equating the computed Q_j to their expectations <u>i.e</u>. by inverting the system (3) to compute the vector $\hat{\sigma}^2$ with elements $\hat{\sigma}_i^2$

$$\hat{\underset{\sim}{\sigma}}^2 = K^- \underset{\sim}{Q}(y) \tag{4}$$

from the vector $\underset{\sim}{Q}(y)$ with elements $Q_j(y)$ where $K = (k_{ji})$ with rank to be discussed in Section 6 and 7.

(d) Replacing any negative elements of $\hat{\sigma}^2$ by 0, with consequences to be discussed in Section 7.

We now give more details for (a), (b) and (c):

(a) The $Q_j(y)$ will be based on contrasts which do not depend on any elements of α. Accordingly we orthogonalize all U_i matrices on X and construct matrices V_i orthogonal on X as follows: Denote by $u(t,i)$ the tth column vector of U_i and by $x(r)$ the rth column vector of X then the columns $v(t,i)$ of V_i are given by

$$\left.\begin{aligned} &v(t,i) = u(t,i) - \sum_{r=1}^{k} x(r)\,\{x'(r)u(t,i)\} \\ &\text{or} \\ &V_i = U_i - XX'U_i \;. \end{aligned}\right\} \tag{5}$$

We now choose the c+1 quadratic forms $Q_j(y)$ as

$$Q_j(y) = y'V_jV_j'y = (V_j'y)'V_j'y \qquad j = 1, \ldots, c+1 \tag{6}$$

(b) It follows from the method of synthesis (see Hartley (1967), J. N. K. Rao (1968)) that

$$E\,Q_j(y) = \sum_{i=1}^{c+1} k_{ji}\sigma_i^2 \qquad \text{with}$$
$$k_{ji} = \sum_t (V_j' u(t,i))'(V_j' u(t,i)) \tag{7}$$

Now since $v(\tau,j)$ is orthogonal on any $x(\rho)$ (i.e. since $v'(\tau,j)x(\rho) = 0$) we can write the k_{ji} in the alternative form

$$k_{ji} = \sum_t (V_j' v(t,i))'(V_j' v(t,i))$$
$$= \sum_{t\tau} \{v'(\tau,j)\, v(t,i)\}^2 \tag{8}$$

showing that $k_{ij} = k_{ji}$.

An alternative form of k_{ji} is

$$k_{ji} = \mathrm{tr}\{(V_j V_j')(V_i V_i')\} \; . \tag{9}$$

We shall show in Section 6 that the symmetrical matrix $K = (k_{ji})$ will have full rank c+1 if the n x n matrices $V_i V_i'$ are not linearly dependent.

(c) We shall also show in Section 6 that the system of equations

$$\underset{\sim}{Q} = K\hat{\underset{\sim}{\sigma}}^2 \tag{10}$$

is consistent even if the rank of K is degenerate. Solving (10) in the form

$$\hat{\underset{\sim}{\sigma}}^2 = K^- \underset{\sim}{Q} \tag{11}$$

we shall, of course, be particularly interested in the full rank case when $K^- = K^{-1}$.

4. The Computational Load

It may be helpful to give an idea of the computational efficiency of the present method by tabulating the number of products involved in the main operations of the algorithm. To this end we first note simplified versions for the $k_{c+1,i}$: Observing that $U_{c+1} = I$ we have from (5) that $V_{c+1} = I - XX'$ and since $X'X = I$ we find that $V_{c+1}V'_{c+1} = I - XX'$ and finally from (9) that

$$\left.\begin{aligned} k_{c+1,c+1} &= \mathrm{tr}\,(I - XX')(I - XX') = \mathrm{tr}\,(I - XX') \\ &= n - k\ . \end{aligned}\right\} \qquad (12)$$

Similarly we find that

$$\left.\begin{aligned} k_{c+1,i} &= \mathrm{tr}\,\{(I - XX')(V_iV_i')\} \\ &= \mathrm{tr}\,\{V_iV_i' - XX'V_iV_i'\} \\ &= \mathrm{tr}\,V_iV_i'\ . \end{aligned}\right\} \qquad (13)$$

Further we note the form of $V'_{c+1}y$ i.e.

$$V'_{c+1}y = y - XX'y\ . \qquad (14)$$

Defining now the adjoined matrices

$$U = (U_1 \mid \cdots \mid U_c) \qquad V = (V_1 \mid \cdots \mid V_c) \qquad (15)$$

the bulk of the work consists of the formation of the elements of the symmetrical matrix $V'V = V'U = U'V$. The elements of this matrix are assembled in submatrices in accordance with the partition (15) as shown in the Schedule 1 below where it must be remembered that the range of the column index t depends on i and is $t = 1, \ldots, m_i$ and the range of $\tau = 1, \ldots, m_j$ so that the submatrix $V_j'U_i$ has dimensions $m_j \times m_i$. The k_{ji} for $i \geq j = 1, \ldots, c$ are then obtained by forming the sums of squares of the elements in each submatrix in accordance with (7).

Finally, we recite the formulas for the remaining coefficients in the equations (10). The $k_{c+1,c+1}$ and $k_{c+1,i}$ are computed from (12) and

Schedule 1: Submatrices of V'U

	U_1	U_2	. . .	U_c
V_1	$v(\tau,1)'u(t,1)$	$v(\tau,1)'u(t,2)$	. . .	$v_1(\tau,1)'u(t,c)$
V_2		$v(\tau,2)'u(t,2)$	. . .	$v(\tau,2)'u(t,c)$
⋮				. . . .
V_c				$v(\tau,c)'u(t,c)$

(13) respectively and the right hand sides of $Q_j(y)$ from the second form in (6) for $j = 1, \ldots, c$ while $Q_{c+1}(y)$ is given in accordance with (14) by

$$Q_{c+1}(y) = y'y - (X'y)'(X'y) \qquad (16)$$

We can now summarize the approximate number of products involved in the various operations of the algorithms.

Operation	Approximate No. of Products Involved
Orthogonalization of X'X	$(1/2)k^+(k^+ - 1)n$ where k^+ = # of columns in original X
$X'U_i \quad i = 1, \ldots, c$	kmn
$X(X'U_i) \quad i = 1, \ldots, c$ (equation (5))	nmk
$U'V = V'V$ (Schedule 1)	0 Subtotals of elements of $v(t,i)$
$k_{ij} \quad i,j = 1, \ldots, c$ (equation (7))	$(1/2)m(m+1)$
$k_{c+1,i} \quad i = 1, \ldots, c$ (equation (13))	mn

Operation	Approximate No. of Products Involved
$k_{c+1,c+1}$ (equation (12))	0
$Q_j(y)$ $j = 1, \ldots, c+1$ (equations (6), 2nd form and (16))	$(m+k+1)(n+1)$

The important point is that the number of products is only a linear function of the number of data lines n. An approximate formula for the total number of products is $n\{\frac{1}{2} k^+(k^+ - 1) + (2m+1)(k+1)\}$

5. A Numerical Example

A small numerical example with $n = 4$, $k^+ = 3$, $k = 2$, $c = 1$, $m_1 = 2$, $m = 2$, $m_2 = n = 4$ is shown in schedule 2 below.

Schedule 2: A Numerical Example of a Mixed Model

y	X Original			U_1		U_2				X new		V_1	
4	1	1	0	1	0	1	0	0	0	(1/2)	(1/2)	+(1/2)	-(1/2)
2	1	1	0	0	1	0	1	0	0	(1/2)	(1/2)	-(1/2)	+(1/2)
1	1	0	1	0	1	0	0	1	0	(1/2)	-(1/2)	0	0
2	1	0	1	0	1	0	0	0	1	(1/2)	-(1/2)	0	0

The orthogonolization of X (original) to X (new) follows the standard Gram Schmidt procedure and reduces the $k^+ = 3$ dependent columns to $k = 2$ columns which are orthogonal and standardized. Note that

$$x(2)_{new} = x(2)_{old} - (1/2)x(1)_{old} \text{ and}$$

$$x(3)_{old} = x(1)_{new} - x(2)_{new} \text{ must be eliminated.}$$

Using now $x(r) = x(r)_{new}$ we orthogonalize U_1 on X and compute (see (5))

$$x'(1)\ u(1,1) = +(1/2),\ x'(2)\ u(1,1) = +(1/2)$$

and hence

$$v(1,1) = u(1,1) - (1/2)x(1) - (1/2)x(2)$$

likewise

$$x'(1)\ u(2,1) = (3/2),\ x'(2)\ u(2,1) = -(1/2)$$

and hence

$$v(1,2) = u(2,1) - (3/2)x(1) + (1/2)x(2)\ .$$

This yields the matrix V_1 in schedule 2 which has only one independent column. The elements of $V_1'U_1$ require the computation of

$$v(1,1)'\ u(1,1) = (1/2);\ v(1,1)'\ u(2,1) = v(2,1)'\ u(1,1) = -(1/2)$$

and

$v(2,1)'\ u(2,1) = 1/2$ with sum of squares of $k_{11} = 4(1/2)^2 = 1$. Further (equation (12)) $k_{22} = 4 - 2 = 2$ and (equation (13)) $k_{12} = k_{21} = 4(1/2)^2 + 4(0)^2 = 1$ so that the K matrix is given by $K = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

Finally, (equation (16))

$$Q_2(y) = 4^2 + 2^2 + 1^2 + 2^2 - (\tfrac{1}{2}\ 9)^2 - (\tfrac{1}{2}\ 3)^2 = 25 - \frac{90}{4} = 25 - 22.5 = 2.5$$

and (equation (6)) $Q_1(y) = (\frac{1}{2}\ 2)^2 + (\frac{1}{2}(-2))^2 = 2$.

The solution of $\underset{\sim}{Q} = K\hat{\underset{\sim}{\sigma}}^2$ therefore yields $\hat{\sigma}_2^2 = 1/2$, $\hat{\sigma}_1^2 = 1.5$.

6. Optimality Properties and the Consistency of the Equations

The estimators described in Section 3 may be seen to be "best at $\sigma_i^2 = 0$, $i = 1, \ldots, c$, $\sigma_{c+1}^2 = 1$" as defined by L. R. LaMotte (1973). Therefore, the consistency of equation (10), regardless of the rank of K, is established as Lemma 4 by LaMotte (1973). That the estimators defined by (11) are "best" among invariant quadratic unbiased estimators guarantees that they are admissible in that class: that is, no other invariant quadratic unbiased estimators have uniformly less variance for all $\underset{\sim}{\sigma}$. Further, as noted by LaMotte (1973), the estimators (11) have the property that in any model for which a uniformly best estimator exists, (11) will be uniformly best. Finally, it may be seen that the "synthesis" estimators (11) are also MINQUE as in Rao (1971, Section 6) with $V = I$. No claim is made that this choice of the norm has any particular merits among the rather general family of the norms covered by Minque formulas. However, it appears to be reasonable to us that in the absence of any theoretical criteria for selection of Minque norms a norm leading to simple estimators may be regarded as meritorious.

Following Section A5 in LaMotte (1973), it may be seen that the rank of K is equal to the number of linearly independent matrices among $V_i V_i'$, $i = 1, \ldots, c+1$. Thus a singular K may occur if the $U_i U_i'$ matrices are not all linearly independent or if there exists (see (5)) a linear combination of the $U_i U_i'$ matrices whose columns are contained in the linear subspace spanned by the columns of X. In the first case the singularity is caused by the design leading to the U_i matrices, while in the second the singularity is caused by confounding fixed and random effects. In either case, (10) is consistent but some linear combinations of the variance components can not then be unbiasedly estimated. We should stress however that other special cases of Minque (not necessarily invariant to α) may also deserve particular attention.

APPENDIX

The Asymptotic Consistency of $\hat{\sigma}^2$

In discussing the asymptotic behavior of $\hat{\sigma}^2$ it is of course necessary to specify the limiting process under which such properties are supposed to hold. Clearly it is necessary for the consistent estimation of the variances $\sigma_i^2 =$ Var b_i that the number of elements m_i in the vectors b_i all tend to ∞. For the identity matrix U_{c+1} we have $m_{c+1} = n$ the overal sample size. For the remaining m_i we assume that their limiting behavior is related to n by

$$Ln^{1-\alpha_i} \leq m_i \leq Un^{1-\alpha_i} \tag{17}$$

where $0 \leq \alpha_i < 1$ and L, U are universal constants. More specifically we assume that $\alpha_{c+1} = 0$ but $\alpha_i > 0$ for $i = 1, \ldots, c$. Generalizations to situations in which $\alpha_i = 0$ for several components are under consideration.
Denote now by

$$\nu(t, i) = \text{number of elements in } u(t, i) \text{ which are } 1 \tag{18}$$

and

$$\nu(t, i; \tau, j) = \text{number of rows in which both } u(t, i) \text{ and } u(\tau, j) \text{ have elements } 1. \tag{19}$$

Using these concepts we introduce the following conditions of 'pseudo orthogonality' of the u(t, i) vectors. We assume that

$$\ell n^{\alpha_i} \leq \nu(t, i) \leq u\, n^{\alpha_i} \tag{20}$$

(where ℓ, u are universal constants) and that

$$\begin{aligned} &\nu(t,\ i;\ \tau,\ j) = o(\nu(t,\ j)) \\ &i \neq j \text{ with } i = 1,\ \ldots,\ c+1 \\ &\text{and } j = 1,\ \ldots,\ c \end{aligned} \tag{21}$$

The relationship between (17) and (20) is obvious since $\sum_{t=1}^{m_i} \nu(t, i) = n$ so that (20) implies (17) with $U = \frac{1}{\ell}$ and $L = \frac{1}{u}$ and the stronger condition (20) implies a uniform order of magnitude for all $\nu(t, i)$ in a given U_i. Since the columns of the U_i matrices are orthogonal we have $\nu(t, i; \tau, i) = 0$ for all pairs $t \neq \tau$. For columns $u(t, i)$, $u(\tau, j)$ with $i \neq j$ condition (21) is satisfied if there is an asymptotically uniform distribution of the $\nu(t, i)$ rows for which $u(t, i)$ has elements 1 over a fraction qm_j of the m_j columns of U_j where $0 < q < 1$ since the fraction of $\nu(t, i)$ which gives rise to $\nu(t, i; \tau, j)$ will be $O(q^{-1} m_j^{-1}) = O(n^{\alpha_j - 1})$ and will tend to zero.

Next we must introduce conditions on the orthogonal standardized matrix X with elements x_{sr}. Denote by $\sum_{s(t,i)} x_{sr}^2$ the sum of x_{sr}^2 over those rows for which $u(t, i)$ has a 1 element then we assume that

$$\sum_{s(t,i)} x_{sr}^2 = O(n^{\alpha_i - 1}) \tag{23}$$

Since $\sum_s x_{sr}^2 = 1$ and the number of terms in $\sum_{s(t,i)}$ is $\nu(t, i) = O(n^{\alpha_i})$ condition (23) implies that asymptotically the x_{sr}^2 have a uniform density $x_{sr}^2 = O(n^{-1})$.

Finally we place on record a consequence of conditions (18) to (23): it follows from (5) using (18), (19), (23) and Schwartz' inequality that

$$u'(t, i)\ v(\tau, j) = \begin{cases} \nu(t, i) + 0(n^{2\alpha_i - 1}) \text{ for } t = \tau,\ i = j \\ 0 + 0(n^{2\alpha_i - 1}) \text{ for } t \neq \tau,\ i = j \\ \nu(t, i; \tau, j) + 0(n^{\alpha_i + \alpha_j - 1}) \text{ for } i \neq j \end{cases} \tag{24}$$

We now turn to the asymptotic behavior of the k_{ii} and k_{ij}. From (8), (17), (20), and (25) we have that

$$k_{ii} = \sum_{t=1}^{m_i} \sum_{\tau=1}^{m_j} (u'(t, i)\ v(\tau, i))^2$$

$$= \sum_{t=1}^{m_i} \left\{u'(t, i)\ v(t, i)\right\}^2 + \sum_{t \neq \tau}^{m_i} \left\{u'(t, i)\ v(\tau, i)\right\}^2 \tag{25}$$

$$\geq \text{Const } n^{1-\alpha_i+2\alpha_i} + 0(n^{2-2\alpha_i+4\alpha_i-2})$$

$$\geq C\ n^{1+\alpha_i} \qquad \text{for all } i = 1, \ldots, c + 1$$

From (8), (17), (19), (21) and (24) we have for $i \neq j$; $i = 1, \ldots, c+1$; $j = 1, \ldots, c$

$$k_{ij} = \sum_{t=1}^{m_i} \sum_{\tau=1}^{m_j} \left\{u'(t, i)\ v(\tau, j)\right\}^2$$

$$= \sum_{t}^{m_i} \sum_{\tau}^{m_j} \nu(t, i; \tau, j)^2 + 0(n^{\alpha_i+\alpha_j-1}) \sum_{t}^{m_i} \sum_{\tau}^{m_j} \nu(t, i; \tau, j)$$

$$+ \sum_{t}^{m_i} \sum_{\tau}^{m_j} 0(n^{2\alpha_i+2\alpha_j-2})$$

$$= \sum_{t}^{m_i} o(\nu(t, i)) \sum_{\tau}^{m_j} \nu(t, i; \tau, j) + 0(n^{\alpha_i+\alpha_j-1})n \tag{26}$$

$$+ 0(n^{2-\alpha_i-\alpha_j})\ 0(n^{2\alpha_i+2\alpha_j-2})$$

$$= o(n^{1+\alpha_i}) + O(n^{\alpha_i+\alpha_j}) = o(n^{1+\alpha_i}) \tag{26}$$

since $\alpha_j < 1$. Similarly we prove by symmetry that $k_{ij} = o(n^{1+\alpha_j})$ for $i \neq j \leq c$. From (25) and (26) it is clear that for all large n the $c \times c$ matrix k_{ij} for $i, j = 1, \ldots, c$ is asymptotically diagonal with diagonal coefficients $\geq cn^{1+\alpha_i}$ while the coefficients $k_{c+1,j}$ are asymptotically equal to $o(n)$. Moreover it is obvious from (12) that $k_{c+1,c+1} \geq Cn$. Using therefore the first c equations of $K\hat{\sigma}^2 = Q(y)$ we obtain that

$$\hat{\sigma}_i^2 = O(n^{-\alpha_i-1}) \{Q_i(y) - o(n)\hat{\sigma}^2_{c+1}\} = O(n^{-\alpha_i-1})Q_i(y) + o(n^{-\alpha_i})\hat{\sigma}^2_{c+1}$$

$$\text{for } i = 1, \ldots, c \tag{27}$$

Substituting (27) in the last equation we obtain

$$\hat{\sigma}^2_{c+1} \{cn + o(n^{1-\alpha_{min}})\} = Q_{c+1}(y) + \sum_{i=1}^{c} Q_i(y)\, o(n^{-\alpha_i}) \tag{28}$$

or

$$\hat{\sigma}^2_{c+1} = O(n^{-1})Q_{c+1}(y) + \sum_{i=1}^{c} Q_i(y)\, o(n^{-\alpha_i-1}) \tag{29}$$

Substituting (29) back in (27) we obtain

$$\hat{\sigma}_i^2 = O(n^{-\alpha_i-1})\, Q_i(y) + o(n^{-1-\alpha_i})\, Q_{c+1}(y) \tag{30}$$

Equations (29) and (30) show that $\hat{\sigma}^2$ is estimable from the $Q_i(y)$. They also show that $\hat{\sigma}^2$ is consistent provided we can show that

$$\begin{aligned} \text{Var } Q_r(y) &= o(n^{2\alpha_r+2}) \\ & \qquad\qquad \text{for } r = 1, \ldots, c \qquad (31) \\ \text{Var } Q_{c+1}(y) &= o(n^2) \end{aligned}$$

since $\text{Cov}Q_i(y)Q_j(y) = O(\text{Var}Q_i(y)^{\frac{1}{2}}\ \text{Var}Q_j(y)^{\frac{1}{2}})$.

In order to prove the first result in (31) we use formulas [22], [32], [33] and [34] of J.N.K. Rao (1968) with slightly altered notation. Formula [22] gives $E\ Q_r^2(y)$ in the form

$$E(Q_r(y)^2) = 2 \sum_{i<j=1}^{c+1}\sum c_{ij}\sigma_i^2\sigma_j^2 + \sum_{i=1}^{c+1} c_{ii}\sigma_i^4 + \sum_{i=1}^{c+1} h_i\mu_{4i} \qquad (32)$$

where $\mu_{4i} = E\ b_{i\ell}^4$ are the 4^{th} moments of the elements $b_{i\ell}$ of b_i. Noting that $\text{Var } Q_r(y) = E\ Q_r(y)^2 - E^2(Q_r(y))$ the leading terms of c_{ii} and c_{ij} given by J.N.K. Rao's *equations* [33] *and* [32] *cancel and we are left to consider the orders* of magnitude of

$$\begin{aligned} \dot{c}_{ii}-2h_i &= \sum_{t<\tau=1}^{m_i}\sum \{Q_r(u(t,\ i) + u(\tau,\ i)) - Q_r(u(t,\ i)) - Q_r(u(\tau,\ i)\}^2 \\ &= \sum_{t<\tau=1}^{m_i}\sum \{ \sum_{s=1}^{m_r} 2(u(t,\ i)'\ v(s,\ r))\ (u(\tau,\ i)'\ v(s,\ r))\}^2 \ . \end{aligned} \qquad (33)$$

Consider first the case $r = i$. We distinguish two terms when $s = t$ and $s = \tau$. For those two terms $(u(t,\ i)'\ v(s,\ i))\ (u(\tau,\ i)'\ v(s,\ i))$ is from (24) of the order of magnitude $O(n^{\alpha_i})\ O(n^{2\alpha_i-1}) = O(n^{3\alpha_i-1})$. For the remaining terms in $\sum_{s=1}^{m_r}$ the product is of the order $O(n^{4\alpha_i-2})$ but the number of terms is of the order

$O(n^{1-\alpha_i})$ so that $\{\sum_s\}^2$ is $O(n^{6\alpha_i-2})$ and hence $\dot{c}_{ii} = O(n^{2-2\alpha_i})\, O(n^{6\alpha_i-2}) = O(n^{4\alpha_i}) = o(n^{2\alpha_i+2})$ since $\alpha_i < 1$.

Consider next the case $r \neq i$ and $r \neq c + 1$. We have from (33) and (24)

$$\begin{aligned}
\dot{c}_{ii} &= \sum_{t<\tau}^{m_i}\sum \{\sum_s (\nu(t, i; s, r)\, \nu(t, i; s, r) + O(n^{2\alpha_i+2\alpha_r-2}) \\
&\quad + O(n^{\alpha_i+\alpha_r-1})\, (\nu(t, i; s, t) + \nu(\tau, i; s, r))\}^2 \\
&= \sum_{t<\tau}^{m_i}\sum \{o(\nu(s, r)) \sum_s \nu(\tau, i; s, r) + O(n^{2\alpha_i+\alpha_r-1}) \\
&\quad + O(n^{\alpha_i+\alpha_r-1})\, (\nu(t, i) + \nu(\tau, i))\}^2 \\
&= \sum_{t<\tau}^{m_i}\sum \{o(n^{\alpha_r+\alpha_i}) + O(n^{2\alpha_i+\alpha_r-1})\}^2 \\
&= o(n^{2+2\alpha_r}) + o(n^{\alpha_i+2\alpha_r+1}) + O(n^{2\alpha_i+2\alpha_r}) \\
&= o(n^{2+2\alpha_r}).
\end{aligned} \tag{34}$$

The case $r \neq i$, $r = c + 1$ follows on the same lines as (34) except that $\alpha_r = 0$ and that $\nu(t, i; s, c+1)\, \nu(\tau, i; s, c+1) = 0$ since $u(s, r)$ has a 1 only in the s^{th} row and either $u(t, i)$ or $u(\tau, i)$ have a zero in that row. The order of magnitude of $\{\}$ will therefore be $O(n^{2\alpha_i-1})$ and $\dot{c}_{ii}$ will be $O(n^{2\alpha_i}) = o(n^2)$.

The treatment of the c_{ij} in J.N.K. Rao's formula [33] follows on similar lines to the above proof for the c_{ii} if of the two alternatives $i < j$, $j < i$ in (21) the smaller α_i, α_j is selected for majorisations.

It remains to consider the terms

$$h_i = \sum_{t=1}^{m_i} Q_r^2(u(t,\ i)) = \sum_{t=1}^{m_i} \{ \sum_{s=1}^{m_r} (u'(t,\ i)\ v(s,\ r))^2\}^2 \tag{35}$$

For the case r = i we have using (24)

$$h_i = \sum_{t=1}^{m_i} \{(u'(t,\ i)\ v(t,\ i))^2 + \sum_{s\neq t}^{m_i} (u'(t,\ i)\ v(s,\ i))^2\}^2$$

$$= \sum_{t=1}^{m_i} \{O(n^{2\alpha_i}) + O(n^{3\alpha_i-1})\}^2 \tag{36}$$

$$= O(n^{1+3\alpha_i}) + O(n^{4\alpha_i}) + O(n^{5\alpha_i-1})$$

$$= o(n^{2\alpha_i+2}) = o(n^{2\alpha_r+2}) \qquad \text{for } i = r \neq c+1,$$

$$= o(n^2) \qquad \text{for } i = r = c+1.$$

For the case i ≠ r and r ≠ c + 1

$$h_i = \sum_{t=1}^{m_i} \{ \sum_{s=1}^{m_r} (\nu(t,\ i;\ s,\ r) + O(n^{\alpha_i+\alpha_r-1}))^2\}^2$$

$$= \sum_{t=1}^{m_i} \{ \sum_{s=1}^{m_r} o(\nu(s,\ r))\nu(t,\ i;\ s,\ r) + O(n^{\alpha_i+\alpha_r-1})\sum_s \nu(t,\ i;\ s,\ r)$$

$$+ O(n^{1-\alpha_r})\ O(n^{2\alpha_i+2\alpha_r-2})\}^2 \tag{37}$$

$$= \sum_{t=1}^{m_i} \{o(n^{\alpha_i+\alpha_r}) + O(n^{2\alpha_i+\alpha_r-1})\}^2$$

$$= o(n^{\alpha_i+2\alpha_r+1}) + o(n^{2\alpha_i+2\alpha_r}) + O(n^{3\alpha_i+2\alpha_r-1})$$

$$= o(n^{2+2\alpha_r}). \tag{37}$$

Finally for $r = c + 1$, $i \neq r$ we have

$$h_i = \sum_{t=1}^{m_i} \{ \sum_{s=1}^{n} (\nu(t, i; s, r) + 0(n^{\alpha_i - 1}))^2\}^2$$

$$= \sum_{t=1}^{m_i} \{ \sum_{s} \nu(t, i; s, r)^2 + \sum_{s} \nu(t, i; s, r)\, 0(n^{\alpha_i - 1}) + 0(n^{2\alpha_i - 1})\}^2 \tag{38}$$

Now since $\nu(t, i; s, c + 1)$ is either 0 or 1 we have that $\sum_s \nu(t, i; s, c + 1)^2 = \sum_s \nu(t, i; s, c + 1) = \nu(t, i)$ so that

$$h_i = \sum_{t=1}^{m_i} \{0(n^{\alpha_i}) + 0(n^{2\alpha_i - 1})\}^2 \tag{39}$$

$$= 0(n^{1-\alpha_i})\, 0(n^{2\alpha_i})$$

$$= o(n^2).$$

Since $\hat{\sigma}^2$ is unbiassed and Cov $(\hat{\sigma}^2) \to 0$ as $n \to \infty$ it follows that $\hat{\sigma}^2$ is consistent. Moreover if we replace any negative $\hat{\sigma}_i^2$ by 0 the resulting statistic say $\tilde{\sigma}_i^2$ has a smaller mean square error and hence is also consistent.

The consistent estimator $\tilde{\sigma}^2$ may serve as a starting value for the iterative maximum likelihood estimation procedure described by Hemmerle and Hartley (1973). Under certain regularity conditions (not discussed here) one single cycle of the iteration will result in asymptotically efficient estimators of σ^2 and α. If the iteration is carried to convergence solutions of the ML equations are reached. If no ML cycles are performed a consistent estimator $\tilde{\alpha}$ of α can be computed from the generalized least squares (ML) equations.

$$\left.\begin{aligned} \tilde{\alpha} &= (X'H^{-1}X)^{-1}(X'H^{-1}y) \\ \text{where } H &= I_n + \sum_{i=1}^{c} \frac{\tilde{\sigma}_i^2}{\tilde{\sigma}_{c+1}^2} U_i U_i' \end{aligned}\right\} \qquad (40)$$

It has been shown by Hemmerle and Hartley (1973) that (40) can be computed directly from the $U_i U_i'$ and $X'U_i$ matrices without the inversion of the $n \times n$ matrix H using their so called W transformation. In fact the W_o matrix (their equation (19)) is essentially given by the $V_i'V_i$ matrices (see the above Schedule 1) and by the contrasts $V_i'y$ required in the computation of $Q_i(y)$.

The variance covariance matrix of $\tilde{\alpha}$ can likewise be computed through the W transformation.

Acknowledgement

One of us (H.O.H.) wishes to acknowledge support from the Army Research Office.

J.N.K. Rao wishes to acknowledge support from the National Research Council of Canada.

References

Gaylor, D. W., Lucas, H. L. and Anderson, R. L. (1970) Calculation of the Expected Mean Squares by the Abbreviated Doolittle and Square Root Methods, Biometrics, 26, 641-656.

Hartley, H. O. (1967) Expectations Variances and Covariances of Anova Mean Squares by 'Synthesis', Biometrics, 23, 105-114.

Hartley, H. O. and Rao, J. N. K. (1967) Maximum Likelihood Estimation for the Mixed Analysis of Variance Model, Biometrika, 54, 93-108.

Hemmerle, W. J. and Hartley, H. O. (1973) Computing Maximum Likelihood Estimates for the Mixed AOV Model Using the W-Transformation, Technometrics, 15, 819-831.

Henderson, C. R. (1953) Estimation of Variance and Covariance Components, Biometrics 9, 226-252.

LaMotte, L. R. (1973) Quadratic Estimation of Variance Components, Biometrics, 29, 311-330.

Liu, L. and Senturia, J. (1976) "Computation of Minque Variance Components Estimates," Technical Report #408, Department of Statistics, University of Wisconsin.

Rao, C. R. (1971) Estimation of Variance and Covariance Components-Minque Theory, Journal of Multivariate Analysis, 1, 257-275.

Rao, J. N. K. (1968) On Expectations Variances and Covariances of ANOVA Mean Squares by 'Synthesis', Biometrics, 24, 963-978.

Searle, S. R. (1971) Linear Models, John Wiley & Sons, Inc., New York.