Biometrika (1971), **58**, 2, *p.* 279
With 3 *text-figures*
Printed in Great Britain

Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance

BY A. T. JAMES
University of Adelaide
AND
G. N. WILKINSON†
C.S.I.R.O., Adelaide

SUMMARY

A factorization is given of the residual operator for nonorthogonal analysis of variance. It is interpreted geometrically in terms of the critical angles between the subspaces determined by the factors. The factorization determines a recursive procedure for analysis as described by Wilkinson (1970). Canonical components are defined and a method of computing them is given together with formulae for their variances, since these would be required for combining information, as for instance, in the recovery of interblock information.

1. INTRODUCTION

Consider a simple randomized block design with t treatments assigned in b blocks of t plots, and let $\mathbf{B}$ and $\mathbf{T}$ be the operators on a vector of observations which replace the observations by the corresponding block or treatment means, respectively. The residual operator $\mathbf{R}$ which produces, from the vector of observations, the vector of deviations from a least squares fit of the usual additive model comprising block and treatment effects, can be expressed in the factorized form

$$\mathbf{R} = (\mathbf{I}-\mathbf{T})(\mathbf{I}-\mathbf{B}) = (\mathbf{I}-\mathbf{B})(\mathbf{I}-\mathbf{T}), \qquad (1\cdot1)$$

in which $\mathbf{I}$ denotes the identity operator.

Similarly, for a balanced incomplete block design, one can deduce from the *relationship algebra* of the design as given by James (1957) the factorization

$$\mathbf{R} = (\mathbf{I}-\mathbf{B})(\mathbf{I}-e^{-1}\mathbf{T})(\mathbf{I}-\mathbf{B}), \qquad (1\cdot2)$$

where $e = (\lambda t)/(rk)$ is the efficiency factor (Yates, 1936) for the design with parameters b, k, t, r and $\lambda = r(k-1)/(t-1)$.

A general recursive relation for specifying factorizations of the residual operator of the kind illustrated above was derived in an unpublished paper by G. N. Wilkinson and is applicable to generally balanced designs, that is, in which each factor of the corresponding model is characterized by a single efficiency factor. The factorization of the residual operator determines a sequence of sweep operations on the data vector, for instance, $(\mathbf{I}-\mathbf{B})$ and $(\mathbf{I}-e^{-1}\mathbf{T})$ in (1·1) and (1·2) above, in each of which a set of effects are calculated and subtracted from the input vector as described by Wilkinson (1970).

† Now at Rothamsted Experimental Station.

The authors independently obtained generalized factorizations applicable to any experimental design. These derive from a polynomial relation for the residual operator which is stated and proved in this paper. If $\mathbf{R}_0$ is the residual operator corresponding to a fit of a linear model up to but excluding a current model factor and $\mathbf{R}$ is the residual operator corresponding to the extended fit including this factor, and if $\mathbf{M}$ is the projection operator on the subspace spanned by the incidence vectors for the parameters of the current factor, then

$$\mathbf{R} = P(\mathbf{Q})\,\mathbf{R}_0, \tag{1·3}$$

where $\mathbf{Q} = \mathbf{R}_0\mathbf{M}\mathbf{R}_0$ is termed here the *shrinkage operator* for the current factor, and $P(\mathbf{Q})$ is the reduced minimum polynomial of the operator $\mathbf{Q}$, normalized with constant term $\mathbf{I}$; see § 4. The correspondence with the particular relation (1·2) above is

$$\mathbf{R}_0 = \mathbf{I}-\mathbf{B}, \quad \mathbf{Q} = \mathbf{R}_0\mathbf{T}\mathbf{R}_0, \quad P(\mathbf{Q}) = \mathbf{I}-e^{-1}\mathbf{Q}.$$

The derivation and interpretation of the fundamental relation (1·3), in relation to an experimental design, depend on the geometrical interrelations of the vector subspaces defined by the incidence vectors for the factors in the corresponding model for analysis, for instance, the block and treatment subspaces in the examples cited above. The characterizing geometrical properties are summarized by a canonical decomposition theorem for vector spaces, due essentially to Jordan and to Hotelling (1936). In § 2 we give a formulation and proof of the theorem in terms of projection and shrinkage operators.

Parallel results have been given by Mann (1960) who obtains them by analysis of the relationship algebra of a design with two nonorthogonal factors.

The significant implication of the decomposition theorem is that the analysis of variance for an experimental design is characterized by the canonical correlations between the subspaces corresponding to factors of the model. The roots of the polynomial $P(\mathbf{Q})$ in (1·3), which are the distinct nonzero eigenvalues of the operator $\mathbf{Q}$, are termed the *canonical efficiency factors* for the corresponding factor of the model. The complements $(1-e_i)$ of the canonical efficiency factors e_i are the squares of the canonical correlation coefficients between the subspace defined by the current factor and that defined by previous factors in the model.

2. A CANONICAL DECOMPOSITION THEOREM FOR VECTOR SPACES

The following results are needed.

A subspace $\mathscr{U}$ in a vector space $\mathbf{R}^n$, uniquely determines the orthogonal projection operator $\mathbf{E}_{\mathscr{U}}$ upon it. The linear operator $\mathbf{E}_{\mathscr{U}}$ is idempotent, $\mathbf{E}_{\mathscr{U}}^2 = \mathbf{E}_{\mathscr{U}}$ and symmetric, $\mathbf{E}'_{\mathscr{U}} = \mathbf{E}_{\mathscr{U}}$. Linear operators can be considered either as matrices with the vectors upon which they operate written as column vectors, or as linear mappings of vectors, e.g. the averaging operators $\mathbf{B}$ and $\mathbf{T}$ above. Conversely, an idempotent symmetric linear operator $\mathbf{A}$ uniquely determines the subspace $\mathscr{U}$, upon which it projects as its range, i.e. the set of vectors of the form $\mathbf{A}\mathbf{x}$ for $\mathbf{x}\in\mathbf{R}^n$. Hence $\mathbf{A}$ and $\mathscr{U}$ determine each other, $\mathbf{A} \leftrightarrow \mathscr{U}$,

$$\mathscr{R}(\mathbf{A}) = \mathscr{U}, \quad \mathbf{A} = \mathbf{E}_{\mathscr{U}}.$$

The operator $\bar{\mathbf{A}} = \mathbf{I}-\mathbf{A}$, which projects orthogonally on the orthogonal complement, $\mathscr{R}(\mathbf{A})^{\perp}$, of $\mathscr{R}(\mathbf{A})$ determines $\mathscr{U}$ as its kernel, $\mathscr{U} = \mathscr{K}(\bar{\mathbf{A}})$, i.e. the set of vectors $\mathbf{u}\in\mathbf{R}^n$ such that $\bar{\mathbf{A}}\mathbf{u} = \mathbf{0}$.

The range and kernel of a symmetric operator $\mathbf{A}$ are orthogonal and span $\mathbf{R}^n$;

$$\mathbf{R}^n = \mathscr{R}(\mathbf{A})\oplus\mathscr{K}(\mathbf{A}).$$

Consider orthogonal projection operators $\mathbf{A}$ and $\mathbf{B}$ in the vector space $\mathbf{R}^n$ and the related symmetric operators $\mathbf{ABA}$, $\mathbf{BAB}$. We have first

LEMMA 1. *The distinct nonzero eigenvalues of* $\mathbf{ABA}$ *and* $\mathbf{BAB}$ *are the same, with the same respective multiplicities.*

Proofs of this, the subsequent theorem and its corollaries are given in the appendix.

Denote the common nonzero eigenvalues of $\mathbf{ABA}$ and $\mathbf{BAB}$ by $\lambda_1, \ldots, \lambda_r$, and let $\mathscr{U}_i$ and $\mathscr{V}_i$ $(i = 1, \ldots, r)$ denote the corresponding eigenspaces of $\mathbf{ABA}$ and $\mathbf{BAB}$ respectively; that is such that $\mathbf{ABAu} = \lambda_i \mathbf{u}$ for $\mathbf{u} \in \mathscr{U}_i$ and $\mathbf{BABv} = \lambda_i \mathbf{v}$ for $\mathbf{v} \in \mathscr{V}_i$. If $\mathscr{U}_0$ and $\mathscr{V}_0$ denote the orthogonal complements of $\mathscr{R}(\mathbf{ABA})$ and $\mathscr{R}(\mathbf{BAB})$ in $\mathscr{R}(\mathbf{A})$ and $\mathscr{R}(\mathbf{B})$ respectively, then

$$\begin{aligned} \mathscr{R}(\mathbf{A}) &= \mathscr{U}_0 \oplus \mathscr{R}(\mathbf{ABA}) = \mathscr{U}_0 \oplus \mathscr{U}_1 \oplus \ldots \oplus \mathscr{U}_r, \\ \mathscr{R}(\mathbf{B}) &= \mathscr{V}_0 \oplus \mathscr{R}(\mathbf{BAB}) = \mathscr{V}_0 \oplus \mathscr{V}_1 \oplus \ldots \oplus \mathscr{V}_r. \end{aligned} \tag{2·1}$$

The relationship between the two subspaces $\mathscr{R}(\mathbf{A})$ and $\mathscr{R}(\mathbf{B})$ is summarized by the following decomposition theorem.

THEOREM 1. *For any two subspaces* $\mathscr{U}$ *and* $\mathscr{V}$ *in* $\mathbf{R}^n$, *let* $\mathbf{A}$ *and* $\mathbf{B}$ *be the respective orthogonal projection operators on them. Then with* $\mathscr{U}_0, \mathscr{V}_0, \mathscr{U}_i$ *and* $\mathscr{V}_i$ *defined as above, the sum* $\mathscr{U} + \mathscr{V}$ *decomposes as follows*:

$$\mathscr{U} + \mathscr{V} = \mathscr{R}(\mathbf{A}) + \mathscr{R}(\mathbf{B}) = \mathscr{U}_0 \oplus \mathscr{V}_0 \oplus (\mathscr{U}_1 + \mathscr{V}_1) \oplus \ldots \oplus (\mathscr{U}_r + \mathscr{V}_r) \tag{2·2}$$

and, for $i = 1, \ldots, r$,

$$\mathbf{A}\mathscr{V}_i = \mathscr{U}_i \quad (\mathbf{A}\mathscr{V}_0 = 0), \quad \mathbf{B}\mathscr{U}_i = \mathscr{V}_i \quad (\mathbf{B}\mathscr{U}_0 = 0), \quad \dim \mathscr{U}_i = \dim \mathscr{V}_i. \tag{2·3}$$

Further, all vectors $\mathbf{u} \in \mathscr{U}_i$ *make the same angle,* θ_i, *with the subspace* $\mathscr{V}_i$, *and vice versa. The* θ_i *are critical angles given by* $\cos^2 \theta_i = \lambda_i$.

COROLLARY 1. *The nonzero eigenvalues and the corresponding eigenspaces of the operators* $\mathbf{AB}$ *and* $\mathbf{BA}$ *are the same as for* $\mathbf{ABA}$ *and* $\mathbf{BAB}$ *respectively, that is,* $\mathbf{ABu} = \lambda_i \mathbf{u}$ *and* $\mathbf{BAv} = \lambda_i \mathbf{v}$ *for* $\mathbf{u} \in \mathscr{U}_i$, $\mathbf{v} \in \mathscr{V}_i$. *The respective ranges are also the same,*

$$\mathscr{R}(\mathbf{AB}) = \mathscr{R}(\mathbf{ABA}), \quad \mathscr{R}(\mathbf{BA}) = \mathscr{R}(\mathbf{BAB}).$$

COROLLARY 2. *If* $\mathbf{A}_i$ *and* $\mathbf{B}_i$ *are the orthogonal projection operators on* $\mathscr{U}_i$ *and* $\mathscr{V}_i$ *respectively,* $(i = 0, 1, \ldots, r)$, *they satisfy the relations*

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1 + \ldots + \mathbf{A}_r, \quad \mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1 + \ldots + \mathbf{B}_r, \tag{2·4}$$

with $\mathbf{A}_0 \mathbf{B}_0 = \mathbf{0}$, *and for* $i \neq j = 0, 1, \ldots, r$,

$$\mathbf{A}_i \mathbf{A}_j = \mathbf{0}, \quad \mathbf{B}_i \mathbf{B}_j = \mathbf{0}, \quad \mathbf{A}_i \mathbf{B}_j = \mathbf{0}, \quad \mathbf{A}\mathbf{B}_i\mathbf{A} = \lambda_i \mathbf{A}_i, \quad \mathbf{B}\mathbf{A}_i\mathbf{B} = \lambda_i \mathbf{B}_i. \tag{2·5}$$

COROLLARY 3. *Consider the symmetric operator* $\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}$, *where* $\bar{\mathbf{A}} = \mathbf{I} - \mathbf{A}$. *Then*

$$\mathscr{R}(\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}) = \mathscr{V}_0 \oplus \mathscr{W}_1 \oplus \ldots \oplus \mathscr{W}_r, \tag{2·6}$$

where $\mathscr{W}_i = (\mathscr{U}_i + \mathscr{V}_i) \cap \mathscr{U}_i^{\perp}$ $(i = 1, \ldots, r)$ *is the orthogonal complement of* $\mathscr{U}_i$ *in* $\mathscr{U}_i + \mathscr{V}_i$, *and the distinct nonzero eigenvalues of* $\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}$ *are* 1, *if* $\mathscr{V}_0$ *is nonnull and* $1 - \lambda_i$ *if* $\lambda_i \neq 1$ $(i = 1, \ldots, r)$, *with corresponding eigenspaces* $\mathscr{V}_0$ *and* $\mathscr{W}_i$ $(i = 1, \ldots, r)$ *excluding the null space* $\mathscr{W}_i$ *corresponding to* $\lambda_i = 1$.

Note that if the ranges of $\mathbf{A}$ and $\mathbf{B}$ are the spaces spanned respectively by the column vectors of matrices $\mathbf{X}$ and $\mathbf{Y}$, which are deviations of variates from their sample means, the values $\lambda_i = \cos^2 \theta_i$ are the squares of the canonical correlation coefficients of the two sets of variates (Hotelling, 1936).

The canonical correlations depend only on the subspaces $\mathscr{R}(\mathbf{X})$ and $\mathscr{R}(\mathbf{Y})$ spanned by the columns of $\mathbf{X}$ and $\mathbf{Y}$. Moreover, the angular relations between the subspaces are invariant under simultaneous orthogonal transformation of $\mathbf{X}$ and $\mathbf{Y}$, i.e. for any orthogonal matrix $\mathbf{H}$, $\mathbf{X} \to \mathbf{HX}$ and $\mathbf{Y} \to \mathbf{HY}$ imply that $\mathbf{A} \to \mathbf{HAH}'$ and $\mathbf{B} \to \mathbf{HBH}'$, so that $\mathbf{ABA} \to \mathbf{HABAH}'$ and its roots $\lambda_i = \cos^2 \theta_i$ are invariant. The critical angles θ_i are invariant under rotation of $\mathbf{R}^n$.

3. Illustration of the decomposition theorem

Consider a lattice design with n^2 treatments arranged in a square; see Table 2*a* for the case $n = 3$. Two replicates are set out, the columns of treatments being taken as blocks in the first replicate and the rows in the second. Let $\mathbf{B}$ and $\mathbf{T}$ denote the orthogonal projection (averaging) operators for blocks and treatments respectively. The pseudo-factorial structure of the treatment grouping, namely rows × columns, suggests the eigenspaces for the operator $\mathbf{TBT}$ given in Table 1.

Table 1. *Eigenspaces of the operator* $\mathbf{TBT}$ *associated with treatments in a lattice design*

Eigenspace of $\mathbf{TBT}$	Dimension	Eigenvalue of $\mathbf{TBT}$	Efficiency factor
Mean	1	1	0
Pseudo-main effects	$2(n-1)$	$\frac{1}{2}$	$\frac{1}{2}$
Pseudo-interactions	$(n-1)^2$	0	1

Table 2. *Transformation of eigenvectors associated with treatments in a lattice design*

(*a*)

Treatments

$$\begin{bmatrix} T_1 & T_2 & T_3 \\ T_4 & T_5 & T_6 \\ T_7 & T_8 & T_9 \end{bmatrix}$$

Design

$$\underset{B_1 \quad B_2 \quad B_3}{\begin{bmatrix} T_1 & T_2 & T_3 \\ T_4 & T_5 & T_6 \\ T_7 & T_8 & T_9 \end{bmatrix}} \quad \begin{bmatrix} T_1 & T_2 & T_3 \\ T_4 & T_5 & T_6 \\ T_7 & T_8 & T_9 \end{bmatrix} \begin{matrix} B_4 \\ B_5 \\ B_6 \end{matrix}$$

(*b*)

A pseudo-main-effect contrast

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{t}$$

Corresponding sample-space vector

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{Xt}.$$

(*c*) Transformation of $\mathbf{Xt}$ by $\mathbf{TBT}$

$$\mathbf{Xt} \xrightarrow{\mathbf{T}} \mathbf{Xt} \xrightarrow{\mathbf{B}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\mathbf{T}} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} = \tfrac{1}{2}\mathbf{Xt}.$$

(*d*) Transformation of $\bar{\mathbf{B}}\mathbf{Xt}$ by $\mathbf{Q} = \bar{\mathbf{B}}\mathbf{T}\bar{\mathbf{B}}$

$$\bar{\mathbf{B}}\mathbf{Xt} \xrightarrow{\bar{\mathbf{B}}} \bar{\mathbf{B}}\mathbf{Xt} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\mathbf{T}} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\bar{\mathbf{B}}} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \tfrac{1}{2}\bar{\mathbf{B}}\mathbf{Xt}.$$

Table 2*c* gives the transformation by **TBT** of a particular sample-space vector **Xt** (Table 2*b*) corresponding to a pseudo-main-effect contrast **t**, and shows that **Xt** is an eigenvector of **TBT** with eigenvalue $\frac{1}{2}$. Clearly therefore the subspace corresponding to all main-effect contrasts is a $2(n-1)$-dimensional eigenspace of **TBT**. Similarly one can readily check that the pseudo-interactions define an $(n-1)^2$-dimensional eigenspace with eigenvalue 0. There is also a one-dimensional eigenspace corresponding to the grand mean, with eigenvalue 1.

Corollary 3 of the decomposition theorem is illustrated by projecting the sample space vector **Xt**, corresponding to the pseudo-main-effect contrast **t**, on the subspace

$$\mathscr{W} = \{\mathscr{R}(\mathbf{B})+\mathscr{R}(\mathbf{T})\} \cap \{\mathscr{R}(\mathbf{B})\}^{\perp} \tag{3·1}$$

to give the vector $\bar{\mathbf{B}}\mathbf{Xt}$, where $\bar{\mathbf{B}} = \mathbf{I}-\mathbf{B}$. By Corollary 3 it is an eigenvector of the operator $\mathbf{Q} = \bar{\mathbf{B}}\mathbf{T}\bar{\mathbf{B}}$ with eigenvalue $1-\frac{1}{2} = \frac{1}{2}$ as verified in Table 2*d*. The eigenvalues of **Q** are the efficiency factors for treatments; see § 4. The zero efficiency factor in Table 1 corresponds to the contrast 'mean treatment effect' which is aliased with the grand mean.

The analysis of the lattice design in Table 2 is discussed in detail by Wilkinson (1970).

4. A Polynomial relation for the residual operator

Suppose that a linear model

$$\mathscr{E}^0(\mathbf{y}) = \mathbf{X}_0\boldsymbol{\tau}_0 \tag{4·1}$$

has been fitted by least squares to a vector of observations **y**, and consider the problem of fitting an extended model with an additional model factor $\mathbf{X}\boldsymbol{\tau}$,

$$\mathscr{E}(\mathbf{y}) = \mathbf{X}_0\boldsymbol{\tau}_0+\mathbf{X}\boldsymbol{\tau}, \tag{4·2}$$

where the vectors of expectations are expressed in terms of the parameter vectors $\boldsymbol{\tau}_0$ and $\boldsymbol{\tau}$ and the corresponding incidence matrices $\mathbf{X}_0$ and **X**.

Let $\mathbf{E}_0$ and **M** denote the orthogonal projection operators on the subspaces spanned by the column vectors of $\mathbf{X}_0$ and **X**, respectively, and let $\mathbf{R}_0$ and **R** be the residual projection operators that produce the vectors of deviations from least squares fits of the models (4·1) and (4·2) respectively.

The operator $\mathbf{R}_0$ is given by

$$\mathbf{R}_0 = \mathbf{I}-\mathbf{E}_0, \tag{4·3}$$

and **R** is the orthogonal projection operator on the orthogonal complement, $\{\mathscr{R}(\mathbf{E}_0)+\mathscr{R}(\mathbf{M})\}^{\perp}$, of the subspace spanned by the column vectors of $\mathbf{X}_0$ and **X**. The relation of **R** to $\mathbf{R}_0$ depends on the relationship of the subspaces $\mathscr{R}(\mathbf{E}_0)$ and $\mathscr{R}(\mathbf{M})$, which in general will be nonorthogonal.

To relate **R** to $\mathbf{R}_0$ we introduce the operator

$$\mathbf{Q} = \mathbf{R}_0\mathbf{M}\mathbf{R}_0, \tag{4·4}$$

termed the *shrinkage operator* for the current factor $\mathbf{X}\boldsymbol{\tau}$ of the model (4·2). The reason for this term will become clear from the geometrical interpretations discussed in § 5.

The relevant canonical decomposition of the sample space $\mathbf{R}^n$ corresponding to the model (4·2) is given by the decomposition theorem in § 2, with the operators **A** and **B** of the theorem identified as $\mathbf{E}_0$ and **M**, respectively. Let

$$\mathscr{M}_0 = \mathscr{R}(\mathbf{E}_0) \cap \mathscr{R}_0(\mathbf{M}) \tag{4·5}$$

be the eigenspace with eigenvalue 1 of the operator $\mathbf{E}_0\mathbf{M}\mathbf{E}_0$. This is the space corresponding to aliased contrasts. Suppose otherwise we have r eigenvalues $\lambda_1, \ldots, \lambda_r$ of

$\mathbf{E}_0\mathbf{M}\mathbf{E}_0$ strictly between 0 and 1. Then in the notations of Theorem 1, the sample space $\mathbf{R}^n$ decomposes as

$$\begin{aligned}\mathbf{R}^n &= \mathscr{M}_0 \oplus \mathscr{U}_0 \oplus \mathscr{V}_0 \oplus (\mathscr{U}_1 + \mathscr{V}_1) \oplus \ldots \oplus (\mathscr{U}_r + \mathscr{V}_r) \oplus \mathscr{R}(\mathbf{R}) \\ &= \mathscr{M}_0 \oplus \mathscr{U}_0 \oplus \mathscr{W}_0 \oplus \mathscr{U}_1 \oplus \mathscr{W}_1 \oplus \ldots \oplus \mathscr{U}_r \oplus \mathscr{W}_r \oplus \mathscr{R}(\mathbf{R}). \end{aligned} \tag{4·6}$$

The subspace $\mathscr{M}_0 \oplus \mathscr{U}_0 \oplus \mathscr{W}_0$ is the orthogonal component of the model, and the mutually orthogonal subspaces $\mathscr{U}_i + \mathscr{V}_j = \mathscr{U}_i \oplus \mathscr{W}_i$ $(i = 1, \ldots, r)$ are the sums of the nonorthogonal canonical subspaces $\mathscr{U}_i$ and $\mathscr{V}_i$ which are eigenspaces of $\mathbf{E}_0\mathbf{M}\mathbf{E}_0$ and $\mathbf{M}\mathbf{E}_0\mathbf{M}$ respectively, corresponding to the eigenvalues $\lambda_i = \cos^2\theta_i$.

The subspace $\mathscr{W}_i = (\mathscr{U}_i + \mathscr{V}_i) \cap \mathscr{U}_i^{\perp}$ $(i = 1, \ldots, r)$ are eigenspaces of the shrinkage operator $\mathbf{Q} = \mathbf{R}_0\mathbf{M}\mathbf{R}_0$ with eigenvalues $e_i = 1 - \lambda_i = \sin^2\theta_i$. Also $\mathscr{W}_0 = \mathscr{V}_0$ is an eigenspace of $\mathbf{Q}$, if nonnull, with eigenvalue 1. Note that the terms of the model (4·2) are orthogonal if and only if $\mathbf{E}_0$ and $\mathbf{M}$ commute, implying that $r = 0$ in (4·6).

Consider now the reduced minimum polynomial of $\mathbf{Q}$, that is, the minimum degree polynomial $P(\mathbf{Q})$ such that

$$\mathbf{Q}P(\mathbf{Q}) = \mathbf{0}. \tag{4·7}$$

The roots of the polynomial $P(\mathbf{Q})$ are the nonzero distinct eigenvalues of the operator $\mathbf{Q}$, and thus correspond to the component spaces $\mathscr{W}_i$ of decomposition in (4·6). This suggests that, for extending the fit of the model (4·1) to that of (4·2), an appropriate analysis will be specified by the relation (4·8) in the following

THEOREM 2. *If $\mathbf{E}_0$ and $\mathbf{M}$ are orthogonal projection operators, then the orthogonal projection $\mathbf{R}$ on the residual space $\{\mathscr{R}(\mathbf{E}_0) + \mathscr{R}(\mathbf{M})\}^{\perp}$ is given by*

$$\mathbf{R} = P(\mathbf{Q})\,\mathbf{R}_0, \tag{4·8}$$

where $\mathbf{R}_0 = \mathbf{I} - \mathbf{E}_0$ and $\mathbf{Q} = \mathbf{R}_0\mathbf{M}\mathbf{R}_0$, and $P(\mathbf{Q})$ is the reduced minimum polynomial of $\mathbf{Q}$ normalized with constant term equal to $\mathbf{I}$.

Proof. Since $\mathbf{M} = \mathbf{E}_0\mathbf{M} + \mathbf{R}_0\mathbf{M}$, we have

$$\begin{aligned}\mathscr{R}(\mathbf{E}_0) + \mathscr{R}(\mathbf{M}) &= \mathscr{R}(\mathbf{E}_0) \oplus \mathscr{R}(\mathbf{R}_0\mathbf{M}) \\ &= \mathscr{R}(\mathbf{E}_0) \oplus \mathscr{R}(\mathbf{Q}),\end{aligned} \tag{4·9}$$

by Corollary 1 of Theorem 1.

Hence we must prove that

(i) $P(\mathbf{Q})\,\mathbf{R}_0$ annihilates $\mathscr{R}(\mathbf{E}_0)$ and $\mathscr{R}(\mathbf{Q})$, and also that

(ii) $P(\mathbf{Q})\,\mathbf{R}_0$ acts as an identity operator on

$$\mathscr{R}(\mathbf{R}) = \{\mathscr{R}(\mathbf{E}_0) \oplus \mathscr{R}(\mathbf{Q})\}^{\perp}.$$

On noting that $\mathbf{R}_0$ and $P(\mathbf{Q})$ commute, (i) follows since $\mathbf{R}_0\mathbf{E}_0 = \mathbf{0}$ and $P(\mathbf{Q})\,\mathbf{Q} = \mathbf{0}$. Now

$$P(\mathbf{Q})\,\mathbf{R}_0 = (\mathbf{I} + \text{terms in } \mathbf{Q})\,\mathbf{R}_0.$$

Since $\mathbf{R}_0$ acts as an identity operator on $\mathscr{R}(\mathbf{R})$ whereas $\mathbf{Q}$ annihilates it, the result (ii) follows. Thus $P(\mathbf{Q})\,\mathbf{R}_0$ is the required orthogonal projection operator.

If the degree K of $P(\mathbf{Q})$ is zero, $\mathbf{R} = \mathbf{R}_0$. Otherwise the relation (4·8) gives the following factorizations of the residual operator,

$$\mathbf{R} = \left\{\prod_{i=1}^{K} (\mathbf{I} - e_i^{-1}\mathbf{Q})\right\}\mathbf{R}_0 = \left\{\prod_{i=1}^{K} \mathbf{R}_0(\mathbf{I} - e_i^{-1}\mathbf{M})\right\}\mathbf{R}_0, \tag{4·10}$$

where the e_i $(i = 1, \ldots, K)$ are the roots of the polynomial P. In relation to (4·6) $K = r + 1$ if $\mathscr{V}_0$ is nonnull, otherwise $K = r$. We term K the *order of balance* of the factor $\mathbf{X}\boldsymbol{\tau}$ in the model (4·2) and the e_i are the *canonical efficiency factors*.

The relation (4·10) provides a recursive specification of the complete factorization of the residual operator for a general linear model comprising several terms, and defines the appropriate analysis process as a sequence of sweep operations of the form $(\mathbf{I}-e^{-1}\mathbf{M})$ on the sample vector. A detailed account of this analysis process is given by Wilkinson (1970), who also describes an adaptive method of analysis for determining the reduced minimum polynomials $P(\mathbf{Q})$ and hence the efficiency factors required for the analysis. The geometrical interpretation of the relation (4·10) is discussed in § 5.

We now give an equivalent relation for the matrix of the reduced normal equations. It is readily shown that a modified form $\tilde{\mathbf{A}}$ of the matrix $\mathbf{A} = \mathbf{X}'\mathbf{R}_0\mathbf{X}$ of the reduced normal equations for the estimates of the parameters $\boldsymbol{\tau}$ in (4·2) has the same reduced minimum polynomial as $\mathbf{Q}$, that is

$$\tilde{\mathbf{A}}P(\tilde{\mathbf{A}}) = \mathbf{0}, \tag{4·11}$$

where P is the polynomial of (4·8).

Theorem 3. *Let $\mathbf{R}_0$ and $\mathbf{M}$ denote orthogonal projection operators, with $\mathbf{M} = \mathbf{XCX}'$, where $\mathbf{C}$ is the inverse or a symmetric positive semi-definite effective inverse of $\mathbf{X}'\mathbf{X}$, satisfying the relation $\mathbf{X}'\mathbf{XCX}'\mathbf{X} = \mathbf{X}'\mathbf{X}$. Define $\mathbf{Q} = \mathbf{R}_0\mathbf{MR}_0$ and $\tilde{\mathbf{A}} = \mathbf{C}^{\frac{1}{2}}\mathbf{AC}^{\frac{1}{2}}$, where $\mathbf{A} = \mathbf{X}'\mathbf{R}_0\mathbf{X}$. Then $\mathbf{Q}$ and $\tilde{\mathbf{A}}$ have the same reduced minimum polynomial P, and the same nonzero eigenvalues with the same respective multiplicities.*

Proof. Let

$$\mathbf{Z} = \mathbf{R}_0\mathbf{XC}^{\frac{1}{2}}. \tag{4·12}$$

Then

$$\mathbf{Q} = \mathbf{R}_0\mathbf{MR}_0 = \mathbf{R}_0\mathbf{XCX}'\mathbf{R}_0 = \mathbf{ZZ}', \tag{4·13}$$

$$\tilde{\mathbf{A}} = \mathbf{C}^{\frac{1}{2}}(\mathbf{R}_0\mathbf{X})'(\mathbf{R}_0\mathbf{X})\mathbf{C}^{\frac{1}{2}} = \mathbf{Z}'\mathbf{Z}. \tag{4·14}$$

It follows that $\mathbf{Q}$ and $\tilde{\mathbf{A}}$ will have the same set of distinct nonzero eigenvalues and multiplicities and hence the same reduced minimum polynomials.

Corollary. *The theorem remains valid for $\tilde{\mathbf{A}} = \mathbf{CA}$ or $\tilde{\mathbf{A}} = \mathbf{AC}$.*

The equality of the reduced minimum polynomials of $\mathbf{Q}$ and $\tilde{\mathbf{A}}$ has useful applications. If the reduced minimum polynomial P of $\mathbf{Q}$ is known, an effective inverse of $\tilde{\mathbf{A}}$ can be found as follows.

Lemma. *Substitution of $\tilde{\mathbf{A}}$ for x in $x^{-1}\{1-P(x)\}$ yields an effective inverse of $\tilde{\mathbf{A}}$.*

Proof. Put $q(x) = x^{-1}\{1-P(x)\}$. Then $x - x^2q(x) = xP(x)$ and hence

$$\tilde{\mathbf{A}} - \tilde{\mathbf{A}}^2q(\tilde{\mathbf{A}}) = \tilde{\mathbf{A}}P(\tilde{\mathbf{A}}) = \mathbf{0}, \tag{4.15}$$

i.e. $\tilde{\mathbf{A}}q(\tilde{\mathbf{A}})\tilde{\mathbf{A}} = \tilde{\mathbf{A}}$. Thus $q(\tilde{\mathbf{A}})$ is an effective inverse of $\tilde{\mathbf{A}}$.

In other situations the eigenvalues of $\mathbf{Q}$ can be calculated from $\tilde{\mathbf{A}}$, as illustrated by the following example. Consider the cyclic incomplete block design whose 5 blocks have treatments (1, 2, 3), (2, 3, 4), (3, 4, 5), (4, 5, 1) and (5, 1, 2), respectively. The matrix $\mathbf{A}$ of the reduced normal equations for treatment effects is given by

$$\mathbf{A} = \tfrac{1}{3}\begin{bmatrix} 6 & -2 & -1 & -1 & -2 \\ -2 & 6 & -2 & -1 & -1 \\ -1 & -2 & 6 & -2 & -1 \\ -1 & -1 & -2 & 6 & -2 \\ -2 & -1 & -1 & -2 & 6 \end{bmatrix}, \tag{4·16}$$

in which each column comprises the treatment totals of deviations from the block means of the corresponding column vector of the 15×5 incidence matrix for treatment effects. Thus $\mathbf{A}$ is a symmetric circulant matrix, which we can represent here in the notation

$$\mathbf{A} = \tfrac{1}{3}\mathscr{C}_5(6, -2, -1). \tag{4·17}$$

With $\mathbf{C} = \frac{1}{3}\mathbf{I}$, we readily obtain

$$\begin{aligned} 9\tilde{\mathbf{A}} &= \mathscr{C}_5(6, -2, -1), \\ (9\tilde{\mathbf{A}})^2 &= \mathscr{C}_5(46, -19, -4), \\ (9\tilde{\mathbf{A}})^3 &= \mathscr{C}_5(360, -175, -5). \end{aligned} \tag{4·18}$$

Clearly $(9\tilde{\mathbf{A}})^3$ is a linear function $\lambda(9\tilde{\mathbf{A}})+\mu(9\tilde{\mathbf{A}})^2$ of $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{A}}^2$, and it is easily shown that $\lambda = -55$, $\mu = 15$. The polynomial $P(\tilde{\mathbf{A}})$ is therefore

$$P(\tilde{\mathbf{A}}) = \mathbf{I} - \tfrac{27}{11}\tilde{\mathbf{A}} + \tfrac{81}{55}\tilde{\mathbf{A}}^2, \tag{4·19}$$

with roots

$$e_1, e_2 = \frac{15 \pm \sqrt{5}}{18}, \tag{4·20}$$

which are the required efficiency factors for the treatment factor. The design thus has second order balance, which corresponds to the fact that it is a partially balanced incomplete block design with two associate classes of treatment.

5. Geometrical interpretations

The factorization (4·10), of the residual operator $\mathbf{R}$ determines a sequence of operations on the sample-space vector, which for convenience we express here as the sequence of triplets

$$\mathbf{R} = (\mathbf{R}_0\mathbf{S}_K\mathbf{R}_0)(\mathbf{R}_0\mathbf{S}_{K-1}\mathbf{R}_0)\ldots(\mathbf{R}_0\mathbf{S}_2\mathbf{R}_0)(\mathbf{R}_0\mathbf{S}_1\mathbf{R}_0), \tag{5·1}$$

where $\mathbf{S}_i = \mathbf{I} - e_i^{-1}\mathbf{M} = \mathbf{I} - \sin^{-2}\theta_i\mathbf{M}$. Note that since $\mathbf{R}_0$ is idempotent, the second $\mathbf{R}_0$ operation of one triplet suffices as the first $\mathbf{R}_0$ of the next triplet. The geometrical interpretation of each triplet of operations is explained below.

With reference to the decomposition (4·6) of the sample space, let $\mathbf{Y}_i = \mathbf{E}_i\mathbf{y}$ be the orthogonal projection of the sample-space vector on the subspace

$$\mathscr{U}_i + \mathscr{V}_i = \mathscr{U}_i \oplus \mathscr{W}_i$$

corresponding to the eigenvalue $\lambda_i = \cos^2\theta_i$. From the decomposition theorem it follows that all transforms of $\mathbf{R}_0\mathbf{Y}_i$ by sequences of the operators $\mathbf{R}_0$ and $\mathbf{M}$ take place in the two-dimensional subspace spanned by $\mathbf{R}_0\mathbf{Y}_i$ and $\mathbf{M}\mathbf{R}_0\mathbf{Y}_i$. Hence the transformations may be drawn in a two-dimensional diagram as in Fig. 1, except for $\mathbf{Y}_i$ which may lie outside the plane.

This figure shows how the triplet of operators $\mathbf{R}_0\mathbf{S}_i\mathbf{R}_0$ annihilates, i.e. maps to zero, the corresponding component vector $\mathbf{Y}_i$. First, the operator $\mathbf{R}_0$ subtracts from $\mathbf{Y}_i$ the vector $\mathbf{X}_0\mathbf{t}_{i0}^{(0)}$, which is the component in the ith subspace $\mathscr{U}_i + \mathscr{V}_i$ of the total regression $\mathbf{X}_0\mathbf{t}_0^{(0)}$ on $\mathbf{X}_0$. The sweep $\mathbf{S}_i$ then produces and subtracts the ith component $\mathbf{X}\mathbf{t}_i$ of the partial regression $\mathbf{X}\mathbf{t}$ on $\mathbf{X}$. Notice that this operation requires more than the subtraction of averages as would be done by an operator $\mathbf{I} - \mathbf{M}$. In fact $\mathbf{I} - \mathbf{M}$ would only map the vector $\mathbf{R}_0\mathbf{Y}_i$ into the vector $\mathbf{b}$ in the orthogonal complement $\mathscr{V}_i^{\perp}$ of $\mathscr{V}_i$, instead of into $\mathscr{U}_i$. However, dilating the operator $\mathbf{M}$ by the reciprocal of the efficiency factor $e_i = \sin^2\theta_i$ produces the required mapping, since, from the similarity of the relevant right-angled triangles in Fig. 1,

$$\frac{|\mathbf{b}-\mathbf{a}|}{|\mathbf{a}|} = \sin\theta_i = \frac{|\mathbf{a}|}{|\mathbf{c}-\mathbf{a}|},$$

so that

$$\frac{|\mathbf{c}-\mathbf{a}|}{|\mathbf{b}-\mathbf{a}|} = \operatorname{cosec}^2\theta_i. \tag{5·2}$$

The second $\mathbf{R}_0$ operation then annihilates the vector

$$\mathbf{c} = \mathbf{S}_i\mathbf{R}_0\mathbf{y} = \mathbf{X}_0\{\mathbf{t}_{i0}-\mathbf{t}_{i0}^{(0)}\} \tag{5·3}$$

in $\mathscr{U}_i$, which is the ith component of the difference between the total and partial regressions on $\mathbf{X}_0$.

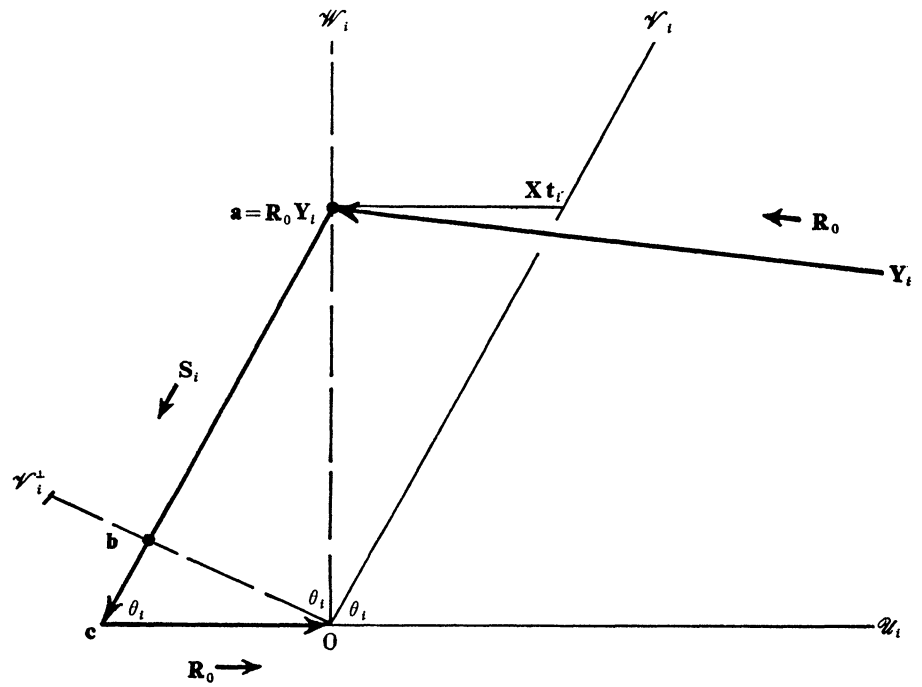

Fig. 1. The transformation by $\mathbf{R}_0\mathbf{S}_i\mathbf{R}_0$ of the orthogonal projection, $\mathbf{Y}_i = \mathbf{E}_i\mathbf{y}$, of $\mathbf{y}$ on $\mathscr{U}_i+\mathscr{V}_i = \mathscr{U}_i\oplus\mathscr{W}_i$; $\mathbf{S}_i = \mathbf{I}-\sin^{-2}\theta_i\mathbf{M}$, $\mathbf{b} = (\mathbf{I}-\mathbf{M})\mathbf{R}_0\mathbf{Y}_i$, $\mathbf{c} = \mathbf{S}_i\mathbf{R}_0\mathbf{Y}_i = \mathbf{X}_0\{\mathbf{t}_{i0}-\mathbf{t}_{i0}^{(0)}\}$.

The efficiency factor $e_i = \sin^2\theta_i$ whose reciprocal e_i^{-1} is needed as a dilation factor can be interpreted as a factor of shrinkage. It is an eigenvalue of $\mathbf{Q} = \mathbf{R}_0\mathbf{M}\mathbf{R}_0$ with eigenspace $\mathscr{W}_i$. The process can again be drawn in a two-dimensional diagram, Fig. 2.

So far, we have considered the effect of the operator $\mathbf{R}_0\mathbf{S}_i\mathbf{R}_0$ acting directly on the component $\mathbf{Y}_i$. When the factorized form (5·1) of the residual operator is applied, the first component $\mathbf{Y}_1$ will be annihilated in this way by the first factor $\mathbf{R}_0\mathbf{S}_1\mathbf{R}_0$, and will remain zero under the operation of the subsequent factors.

Any other component such as $\mathbf{Y}_i$ will first be subjected to $(i-1)$ operators of the factorization (5·1) before the operator $\mathbf{R}_0\mathbf{S}_i\mathbf{R}_0$ is applied. However, the transforms of $\mathbf{Y}_i$ remain within the subspace $\mathscr{U}_i+\mathscr{V}_i$ during the first $(i-1)$ operations and then the operator $\mathbf{R}_0\mathbf{S}_i\mathbf{R}_0$, which annihilates the entire subspace, must map the resultant vector on zero.

In the next section, we make use of the successive transforms of $\mathbf{Y}$ to obtain equations for the canonical components $\mathbf{t}_i$ of $\mathbf{t}$ from which they can be calculated.

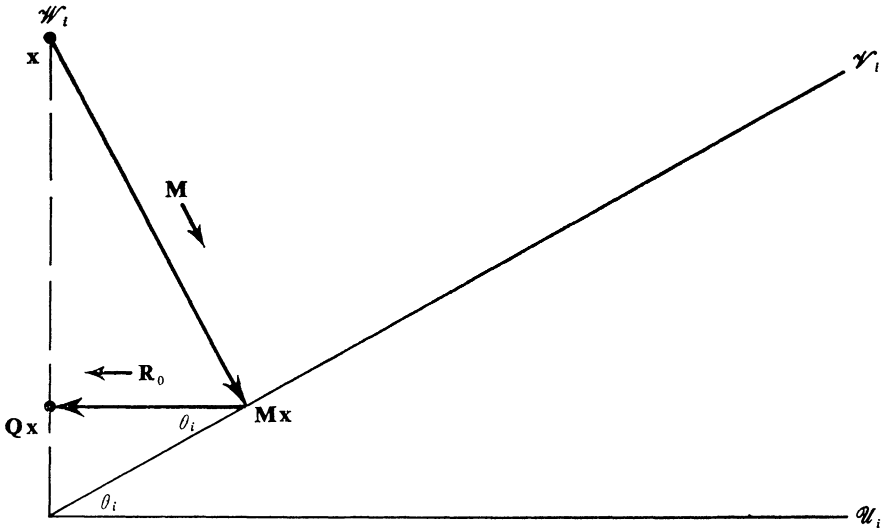

Fig. 2. Transformation by the shrinkage operator $\mathbf{Q}$ of an eigenvector $\mathbf{x}$ in the eigenspace $\mathscr{W}_i$. Here $\mathbf{Q} = \mathbf{R}_0\mathbf{M}\mathbf{R}_0$, $\mathbf{Q}\mathbf{x} = \sin^2\theta_i\mathbf{x}$.

In a multi-stratum analysis, to obtain the analysis of any error stratum other than the lowest, residual stratum, it is necessary to substitute in the sequence (5·1) a pivotal sweep for the appropriate stratum (Wilkinson, 1970). For an incomplete block design, for instance, the pivotal sweep generates a sample space vector comprising estimated block effects, $\mathbf{Y}^{(b)} = \mathbf{E}_0\mathbf{y}$. The mapping operation is illustrated in Fig. 3. The figure also illustrates geometrically the relation between the intrablock and interblock component estimates of treatment effects. The component estimates $\mathbf{t}_i^{(c)}$ of treatment effects ignoring blocks are a weighted combination of the corresponding intra and interblock components,

$$\mathbf{t}_i^{(c)} = \mathbf{t}_i \sin^2\theta_i + \mathbf{t}_i^{(b)} \cos^2\theta_i. \tag{5·4}$$

This weighting is statistically appropriate only when the intra and interblock stratum variances are the same. Otherwise the two components in equation (5·4) are additionally weighted in inverse proportion to the corresponding stratum variances.

Note that Fig. 3 gives only a projected two-dimensional representation of the essential geometry. The vectors $\mathbf{X}\mathbf{t}_i$, $\mathbf{X}\mathbf{t}_i^{(b)}$ which are superimposed collinearly in the figure are in general noncollinear.

6. Calculation of the components of regression in the canonical subspaces

The regression vector $\mathbf{X}\mathbf{t} \in \mathscr{V} = \mathscr{V}_1 \oplus \ldots \oplus \mathscr{V}_K$ must decompose into a sum of its orthogonal projections $\mathbf{E}_{\mathscr{V}_j}\mathbf{X}\mathbf{t}$ on the respective orthogonal subspaces $\mathscr{V}_j$ for $j = 1, \ldots, K$, where $\mathbf{E}_{\mathscr{V}_j}$ is the operator which projects orthogonally on $\mathscr{V}_j$. The vectors $\mathbf{t}_j$ such that

$$\mathbf{E}_{\mathscr{V}_j}\mathbf{X}\mathbf{t} = \mathbf{X}\mathbf{t}_j \quad (j = 1, \ldots, K) \tag{6·1}$$

are called the *canonical components* of regression. The following argument shows how they may be calculated from the successive transforms of $\mathbf{y}$ by factors in (5·1) for the residual operator.

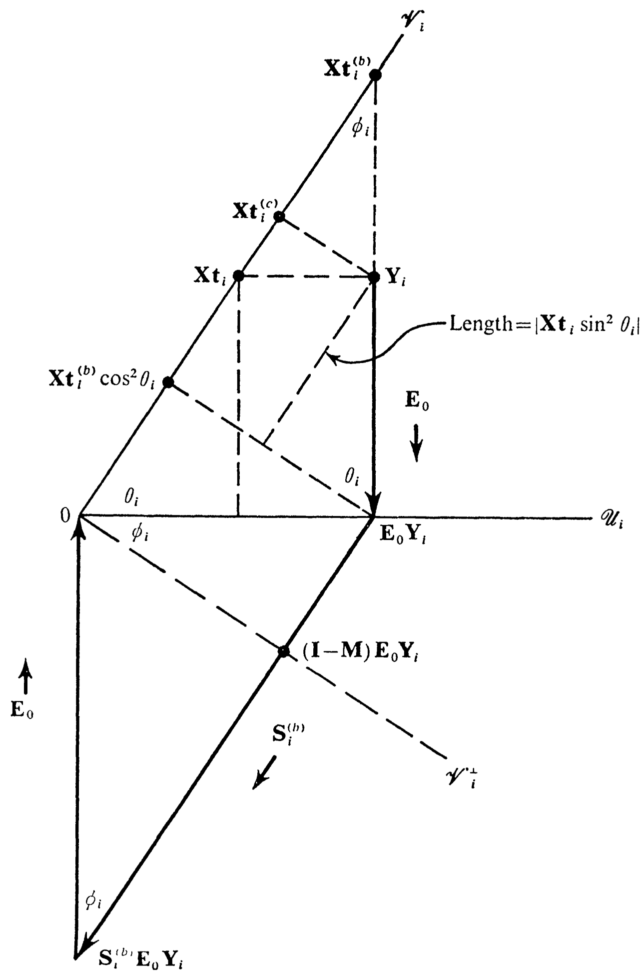

Fig. 3. Interblock analysis for a single treatment factor in a partially balanced incomplete block design. Transformation of $\mathbf{Y}_i = \mathbf{E}_i\mathbf{y}$ by $\mathbf{E}_0\mathbf{S}_i^{(b)}\mathbf{E}_0$; $\mathbf{S}_i^{(b)} = \mathbf{I} - \sin^{-2}\phi_i\mathbf{M} = \mathbf{I} - \cos^{-2}\theta_i\mathbf{M}$, $\mathbf{E}_0$ = pivotal sweep for block factor, $\mathbf{t}_i$ = ith component of intrablock treatment estimates, $\mathbf{t}_i^{(b)}$ = ith component of interblock treatment estimates and $\mathbf{t}_i^{(c)}$ = ith component of treatment estimates ignoring blocks $= \mathbf{t}_i\sin^2\theta_i + \mathbf{t}_i^{(b)}\cos^2\theta_i = \mathbf{t}_i e_i + \mathbf{t}_i^{(b)} e_i^{(b)}$.

Suppose that the operators $\mathbf{S}_1, \ldots, \mathbf{S}_i, \ldots, \mathbf{S}_K$ within the factors $\mathbf{R}_0\mathbf{S}_1\mathbf{R}_0, \ldots, \mathbf{R}_0\mathbf{S}_i\mathbf{R}_0, \ldots, \mathbf{R}_0\mathbf{S}_K\mathbf{R}_0$ successively subtract $\mathbf{X}\tilde{\mathbf{t}}_1, \ldots, \mathbf{X}\tilde{\mathbf{t}}_i, \ldots, \mathbf{X}\tilde{\mathbf{t}}_K$ from $\mathbf{y}$. From them the observed components $\tilde{\mathbf{t}}_1, \ldots, \tilde{\mathbf{t}}_i, \ldots, \tilde{\mathbf{t}}_K$ of treatments can then be found. Put

$$\mathbf{s}_i = \sum_{k=i}^{K} \tilde{\mathbf{t}}_k. \tag{6·2}$$

An equation for $\mathbf{s}_i$ in terms of the canonical components $\mathbf{t}_1, \ldots, \mathbf{t}_i$ of treatments can be deduced as follows. We consider the successive effect of the operators $\mathbf{S}_1, \ldots, \mathbf{S}_{i-1}$ on $\mathbf{Xt}_j$ for $j \geqslant i$.

The middle operator $\mathbf{S}_1$ in the sequence $\mathbf{R}_0\mathbf{S}_1\mathbf{R}_0$, which annihilates the first canonical component $\mathbf{Xt}_1$, also subtracts $e_j e_1^{-1}\mathbf{Xt}_j$ from the jth canonical component $\mathbf{Xt}_j$ leaving $\mathbf{X}(1 - e_j e_1^{-1})\,\mathbf{t}_j$. Likewise, the operator $\mathbf{S}_2$ in the next sequence, $\mathbf{R}_0\mathbf{S}_2\mathbf{R}_0$, which annihilates the rest of the second canonical component $\mathbf{X}(1-e_2e_1^{-1})\mathbf{t}_2$, subtracts $\mathbf{X}e_je_2^{-1}(1-e_je_1^{-1})\mathbf{t}_j$ from the jth canonical component, leaving $\mathbf{X}(1-e_je_2^{-1})(1-e_je_1^{-1})\mathbf{t}_j$. By the time the $(i-1)$ operators have been applied, the vector $\mathbf{t}_j$ will be reduced to

$$\prod_{h=1}^{i-1}(1-e_je_h^{-1})\,\mathbf{t}_j. \tag{6·3}$$

Hence we have equation (6·4) of the following theorem

THEOREM 4. *If* $\mathbf{s}_i$ *is the vector defined in* (6·2), *then the canonical components* $\mathbf{t}_j$ *of* $\mathbf{t}$ *satisfy the equations*

$$\sum_{j=i}^{K}\left\{\prod_{h=1}^{i-1}(1-e_je_h^{-1})\right\}\mathbf{t}_j=\mathbf{s}_i\quad(i=1,\ldots,K) \tag{6·4}$$

whose solution is, for $j=1,\ldots,K$,

$$\mathbf{t}_j=\sum_{k=j}^{K}\left[\left(\prod_{h=1}^{k-1}e_h\right)\left\{\prod_{\substack{h=1\\h\neq j}}^{k}(e_h-e_j)^{-1}\right\}\right]\mathbf{s}_k. \tag{6·5}$$

Proof. If equation (6·5) is used to substitute for $\mathbf{t}_j$ in equation (6·4), then the coefficient of $\mathbf{s}_i$ in the resulting equation for $\mathbf{s}_i$ is 1, and the coefficient of $\mathbf{s}_k$ for $k<i$ is 0. The coefficient of $\mathbf{s}_k$ in the expression for $\mathbf{s}_i$ when $k>i$ is

$$\left(\prod_{h=i}^{k-1}e_h\right)\left\{\sum_{j=i}^{k}\frac{1}{\prod\limits_{\substack{h=i\\h\neq j}}^{k}(e_h-e_j)}\right\}.$$

Since the second factor is symmetric in the indices $i, i+1, \ldots, k$, it must be zero, because if the denominators in it were brought to a common denominator consisting of the difference product

$$\prod_{i\leqslant j_1<j_2\leqslant k}(e_{j_1}-e_{j_2}),$$

which is skew symmetric, the resulting numerator would also have to be skew symmetric, but being of less degree, would thus be zero.

As an example, let us calculate the components $\mathbf{t}_1$ and $\mathbf{t}_2$ of the intrablock estimate of treatment effect for the example given by Wilkinson (1970, Table 3). As we are not going to perform matrix multiplication on our numerical vectors, we take the liberty of leaving them written as arrays.

From his table, we have $K=2$, $e_1=\frac{1}{2}$, $e_2=1$

$$\tilde{\mathbf{t}}_1=\begin{bmatrix}2&1&0\\-1&0&1\\-4&-1&2\end{bmatrix},\quad \tilde{\mathbf{t}}_2=\begin{bmatrix}-1&0&1\\0&0&0\\1&0&-1\end{bmatrix}.$$

Hence

$$\mathbf{s}_1=\mathbf{t}=\tilde{\mathbf{t}}_1+\tilde{\mathbf{t}}_2=\begin{bmatrix}1&1&1\\-1&0&1\\-3&-1&1\end{bmatrix},\quad \mathbf{s}_2=\tilde{\mathbf{t}}_2.$$

From formula (6·5)

$$\mathbf{t}_1=\mathbf{s}_1+\mathbf{s}_2=\begin{bmatrix}0&1&2\\-1&0&1\\-2&-1&0\end{bmatrix},\quad \mathbf{t}_2=-\mathbf{s}_2=\begin{bmatrix}1&0&-1\\0&0&0\\-1&0&1\end{bmatrix}.$$

7. Variance of components and contrasts

Theorem 5. *The variance* $V(\mathbf{Xt}_i)$ *of the ith canonical component* $\mathbf{Xt}_i$ *is given by*

$$V(\mathbf{Xt}_i) = \sigma^2 e_i^{-1} \mathbf{E}_{\mathscr{V}_i},$$

where σ^2 *is the error variance,* e_i *the ith canonical efficiency factor and* $\mathbf{E}_{\mathscr{V}_i}$ *the matrix of the orthogonal projection on the ith canonical subspace* $\mathscr{V}_i$ *of* $\mathscr{R}(\mathbf{X})$.

Proof. Let $\mathbf{X}_i$ be the matrix whose columns are the components of the columns of $\mathbf{X}$ in the subspace $\mathscr{U}_i + \mathscr{V}_i = \mathscr{U}_i \oplus \mathscr{W}_i$. Then

$$\mathbf{X}_i = \mathbf{E}_{\mathscr{U}_i \oplus \mathscr{W}_i} \mathbf{X} = \mathbf{E}_{\mathscr{V}_i} \mathbf{X},$$

and since $\mathbf{Xt}_i$ is a vector in this subspace,

$$\mathbf{Xt}_i = \mathbf{E}_{\mathscr{V}_i} \mathbf{Xt}_i = \mathbf{X}_i \mathbf{t}_i.$$

The partial regression vector $\mathbf{t}_i$ on $\mathbf{X}_i$ is the same as the total regression vector on the orthogonalized vectors $\tilde{\mathbf{X}}_i = \mathbf{R}_0 \mathbf{X}_i$ and since $\tilde{\mathbf{X}}_i \mathbf{t}_i$ is the orthogonal projection of the sample vector $\mathbf{y}$ on $\mathscr{W}_i$, we have $V(\tilde{\mathbf{X}}_i \mathbf{t}_i) = \sigma^2 \mathbf{E}_{\mathscr{W}_i}$.

From the decomposition theorem, or the geometrical interpretation of the shrinkage operator,

$$\mathbf{X}_i \mathbf{t}_i = e_i^{-1} \mathbf{E}_{\mathscr{V}_i} \tilde{\mathbf{X}}_i \mathbf{t}_i.$$

Hence

$$\begin{aligned} V(\mathbf{X}_i \mathbf{t}_i) &= e_i^{-1} \mathbf{E}_{\mathscr{V}_i} \sigma^2 \mathbf{E}_{\mathscr{W}_i} \mathbf{E}_{\mathscr{V}_i} e_i^{-1} \\ &= e_i^{-2} e_i \mathbf{E}_{\mathscr{V}_i} \sigma^2 \\ &= \sigma^2 e_i^{-1} \mathbf{E}_{\mathscr{V}_i}. \end{aligned}$$

Corollary. *For any estimable contrast* $\boldsymbol{\gamma}'\mathbf{t}$,

$$V(\boldsymbol{\gamma}'\mathbf{t}) = V(\mathbf{c}'\mathbf{Xt}) = \sigma^2 \sum_{i=1}^{K} e_i^{-1} \mathbf{c}_i' \mathbf{c}_i,$$

where $\mathbf{c} = \mathbf{XC}\boldsymbol{\gamma}$, $\mathbf{C}$ *is an effective inverse of* $\mathbf{X}'\mathbf{X}$, *and* $\mathbf{c}_i$ *is the component of* $\mathbf{c}$ *in* $\mathscr{V}_i$, *i.e.*

$$\mathbf{c}_i = \mathbf{E}_{\mathscr{V}_i} \mathbf{c}.$$

Proof. Since $\boldsymbol{\gamma} \in \mathscr{R}(\mathbf{X}')$ and $\mathbf{XCX}'\mathbf{X} = \mathbf{X}$,

$$V(\boldsymbol{\gamma}'\mathbf{t}) = V\{(\mathbf{XC}\boldsymbol{\gamma})'\mathbf{Xt}\} = V(\mathbf{c}'\mathbf{Xt}) = \Sigma V(\mathbf{c}_i'\mathbf{Xt}_i) = \Sigma e_i^{-1}\sigma^2 \mathbf{c}_i' \mathbf{E}_{\mathscr{V}_i} \mathbf{c}_i = \sigma^2 \Sigma e_i^{-1} \mathbf{c}_i' \mathbf{c}_i.$$

In order to calculate the components $\mathbf{c}_i$ of $\mathbf{c} = \mathbf{XC}\boldsymbol{\gamma}$ for a given contrast vector $\boldsymbol{\gamma}$, we introduce vectors $\mathbf{t}_i^*$ which are defined modulo the kernel of $\mathbf{X}$, $\mathscr{K}(\mathbf{X})$, by $\mathbf{Xt}_i^* = \mathbf{c}_i$.

Since

$$\mathbf{XC}\boldsymbol{\gamma} = \mathbf{c} = \sum_i \mathbf{c}_i = \mathbf{X} \sum_i \mathbf{t}_i^*$$

their sum $\mathbf{t}^* = \Sigma \mathbf{t}_i^*$ will be given by

$$\mathbf{t}^* = \mathbf{C}\boldsymbol{\gamma} \bmod \mathscr{K}(\mathbf{X}).$$

The components $\mathbf{t}_i^*$ can be obtained from an analysis of $\mathbf{c} = \mathbf{XC}\boldsymbol{\gamma}$ in the same way as the components $\mathbf{t}_i$ of $\mathbf{t}$ were obtained from $\mathbf{y}$. From them, the $\mathbf{c}_i = \mathbf{Xt}_i^*$ can be calculated or the variance of the contrast

$$V(\boldsymbol{\gamma}'\mathbf{t}) = \sigma^2 \sum_{i=1}^{K} e_i^{-1} \mathbf{t}_i^{*\prime} \mathbf{X}'\mathbf{Xt}_i^*.$$

For the example treated in the previous section, suppose we wish to calculate $V(\boldsymbol{\gamma}'\mathbf{t})$ when $\boldsymbol{\gamma}$ is the vector whose elements are given by the array

$$\boldsymbol{\gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The calculations for the analysis are given in Table 3. In this case, $\mathbf{X}'\mathbf{X} = 2\mathbf{I}_9$.

Table 3. *Calculation of the variance of a contrast*

$$\boldsymbol{\gamma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{(\mathbf{X}'\mathbf{X})^{-1}} \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbf{X}} \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbf{R}_0} \frac{1}{6}\begin{bmatrix} 2 & 1 & 0 \\ -1 & -2 & 0 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{\mathbf{t}}_1^* = \frac{1}{6}\begin{bmatrix} 4 & 0 & -1 \\ 0 & -4 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbf{S}_1} \frac{1}{6}\begin{bmatrix} -2 & 1 & 1 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -2 & -1 & 0 \\ 1 & 2 & 0 \\ 1 & -1 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbf{R}_0} \frac{1}{6}\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}, \quad \tilde{\mathbf{t}}_2^* = \frac{1}{6}\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbf{S}_2} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{s}_1^* = \tilde{\mathbf{t}}_1^* + \tilde{\mathbf{t}}_2^* = \frac{1}{6}\begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{s}_2^* = \tilde{\mathbf{t}}_2^*,$$

$$\mathbf{t}_1^* = \mathbf{s}_1^* + \mathbf{s}_2^* = \frac{1}{6}\begin{bmatrix} 2 & 0 & 1 \\ 0 & -2 & -1 \\ 1 & -1 & 0 \end{bmatrix}$$

$$\mathbf{t}_2^* = -\mathbf{s}_2^* = \frac{1}{6}\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{X} = 2\mathbf{I}_9,$$

$$\mathbf{t}_1^{*\prime}\mathbf{X}'\mathbf{X}\mathbf{t}_1^* = \tfrac{2}{3}, \quad \mathbf{t}_2^{*\prime}\mathbf{X}'\mathbf{X}\mathbf{t}_2^* = \tfrac{1}{3},$$

$$V(\boldsymbol{\gamma}'\mathbf{t}) = \sigma^2\left(\frac{1}{\frac{1}{2}}\cdot\frac{2}{3} + \frac{1}{1}\cdot\frac{1}{3}\right) = \tfrac{5}{3}\sigma^2.$$

8. CONCLUSION

Canonical decomposition reveals the structure of an experimental design, in particular the order of balance. The canonical efficiency factors measure the extent of the nonorthogonality. The best methods of computation are thereby indicated, and also the degree of statistical dependence in the estimates, which affects the statistical interpretation. Designs with maximum value for the lowest efficiency factor will usually be advantageous.

The number of degrees of freedom associated with the ith canonical component is the multiplicity of the ith root of the matrix of the reduced normal equations. A simpler method of calculating this would be desirable.

The use of operators provides a much more convenient treatment of the analysis of experimental designs, both for computational as well as for algebraic purposes, than the use of the commonly used multiple-subscript notation for totals and means. The abstract mathematical operators also provide simple computational operators realized as computer subroutines. The addition and multiplication of the operators is given by the relationship algebra (James, 1957) which they constitute.

A thorough treatment of the combination of information from different error strata, as, for example, in the recovery of interblock information, will require decomposition of effects into their canonical components. It will be discussed in a later paper.

Part of the research by A. T. James was carried out at Yale University with support from the U.S. Army, Navy, Airforce and NASA under a contract administered by the Office of Naval Research. Some of the research by G. N. Wilkinson was done at the University of Wisconsin, supported by grants from the General Electric Company, the Graduate School and the Computing Center of the University of Wisconsin.

APPENDIX

Proof of Lemma 1 *and Theorem* 1. Define $\lambda_1, \ldots, \lambda_r$, first, to be the distinct nonzero eigenvalues of **ABA**. Since this operator is symmetric, the corresponding eigenspaces $\mathscr{U}_i$ are orthogonal and

$$\mathscr{R}(\mathbf{ABA}) = \mathscr{U}_1 \oplus \ldots \oplus \mathscr{U}_r. \tag{1}$$

We prove now that the λ_i are also the nonzero eigenvalues of **BAB**. If $\mathbf{u}_i \in \mathscr{U}_i$, then from the relations $\mathbf{B}^2 = \mathbf{B}$, $\mathbf{u}_i = \mathbf{A}\mathbf{u}_i$ and $\mathbf{ABA}\mathbf{u}_i = \lambda\mathbf{u}_i$, we have

$$(\mathbf{BAB})(\mathbf{B}\mathbf{u}_i) = \mathbf{BAB}\mathbf{u}_i = \mathbf{BABA}\mathbf{u}_i = \lambda_i \mathbf{B}\mathbf{u}_i \tag{2}$$

Hence the λ_i are eigenvalues of **BAB**, and the vectors $\mathbf{B}\mathbf{u}_i \in \mathbf{B}\mathscr{U}_i$ are eigenvectors of **BAB**.

Conversely, if λ is any nonzero eigenvalue of **BAB**, and **v** is an associated eigenvector satisfying the relation $\mathbf{BABv} = \lambda\mathbf{v}$, then a similar argument to that in (2) gives

$$\mathbf{ABA}(\mathbf{Av}) = \mathbf{ABAv} = \mathbf{ABABv} = \lambda\mathbf{Av}, \tag{3}$$

so that λ is also an eigenvalue of **ABA**, $\lambda = \lambda_i$ say. Hence the nonzero eigenvalues of **ABA** and **BAB** are the same. Since **Av** is an eigenvector of **ABA** corresponding to the eigenvalue λ_i, it belongs to $\mathscr{U}_i$, i.e. $\mathbf{Av} \in \mathscr{U}_i$.

Furthermore, $\mathbf{BAv} = \mathbf{BABv} = \lambda_i\mathbf{v}$, so that $\mathbf{v} = \lambda_i^{-1}\mathbf{B}(\mathbf{Av}) \in \mathbf{B}\mathscr{U}_i$. Thus $\mathbf{B}\mathscr{U}_i$ is the complete eigenspace $\mathscr{V}_i$ of **BAB** corresponding to λ_i, that is, $\mathscr{V}_i = \mathbf{B}\mathscr{U}_i$. Likewise, from symmetry considerations, $\mathscr{U}_i = \mathbf{A}\mathscr{V}_i$. Since the subspaces $\mathscr{U}_i$ and $\mathscr{V}_i$ map onto each other, their dimensions are the same, i.e. $\dim \mathscr{U}_i = \dim \mathscr{V}_i \; (i = 1, \ldots, r)$.

To prove that $\mathbf{A}\mathscr{V}_0 = 0$ (and likewise, by a similar argument, that $\mathbf{B}\mathscr{U}_0 = 0$) note that if $\mathbf{v}_0 \in \mathscr{V}_0$, then since $\mathbf{v}_0 \in \mathscr{R}(\mathbf{B}) \cap \mathscr{K}(\mathbf{BAB})$ we have $\mathbf{BAv}_0 = \mathbf{BABv}_0 = \mathbf{0}$. Hence $\mathbf{Av}_0 \in \mathscr{K}(\mathbf{B})$ and is therefore orthogonal to $\mathbf{v}_0 \in \mathscr{R}(\mathbf{B})$, so that $\mathbf{0} = \mathbf{v}_0'\mathbf{Av}_0 = (\mathbf{Av}_0)'(\mathbf{Av}_0)$ which implies that $\mathbf{Av}_0 = \mathbf{0}$.

It follows from this that $\mathscr{V}_0$ is orthogonal to $\mathscr{R}(\mathbf{A})$. Likewise $\mathscr{U}_0$ is orthogonal to $\mathscr{R}(\mathbf{B})$. We now prove that $\mathscr{U}_i$ is orthogonal to $\mathscr{V}_j \; (i \neq j;\; i, j = 1, \ldots, r)$.

Let $\mathbf{A}_i$ be the operator which projects orthogonally on $\mathscr{U}_i$. Then, since $\mathscr{R}(\mathbf{A}_i) \subset \mathscr{R}(\mathbf{A})$,

$$\mathbf{A}_i\mathscr{V}_j = \mathbf{A}_i\mathbf{A}\mathscr{V}_j = \mathbf{A}_i\mathscr{U}_j = 0. \tag{4}$$

Hence $\mathscr{V}_j \subset \mathscr{K}(\mathbf{A}_i)$ and is therefore orthogonal to $\mathscr{U}_i = \mathscr{R}(\mathbf{A}_i)$.

This completes the proof for part two of the theorem. Part three is proved as follows:

Consider $\mathbf{u} \in \mathscr{U}_i$ such that $\mathbf{u}'\mathbf{u} = 1$, and let θ_i denote the angle between **u** and its projection **Bu** in $\mathscr{R}(\mathbf{B})$. Then, since $\mathbf{Au} = \mathbf{u}$, $(\cos\theta_i)^2 = \mathbf{u}'\mathbf{Bu} = \mathbf{u}'\mathbf{ABAu} = \lambda_i\mathbf{u}'\mathbf{u}, = \lambda_i$.

Corollary 1 follows straightforwardly from the derivations of the theorem. Corollary 2 is a restatement of the decomposition theorem in terms of the corresponding matrix operators.

The proof of Corollary 3 is as follows:

$$\begin{aligned}\mathscr{R}(\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}) &= \mathscr{R}(\bar{\mathbf{A}}\mathbf{B}) \\ &= \bar{\mathbf{A}}\mathscr{R}(\mathbf{B}) \\ &= \bar{\mathbf{A}}(\mathscr{V}_0 \oplus \mathscr{V}_1 \oplus \ldots \oplus \mathscr{V}_r) \\ &= \mathscr{V}_0 \oplus \mathscr{W}_1 \oplus \ldots \oplus \mathscr{W}_r. \end{aligned} \tag{5}$$

Consider any vector $\mathbf{w}_i \in \mathscr{W}_i$. Then $\mathbf{w}_i$ is of the form $\mathbf{w}_i = \bar{\mathbf{A}}\mathbf{v}_i$, where $\mathbf{v}_i \in \mathscr{V}_i$. Hence

$$\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}\mathbf{w}_i = \bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}\mathbf{v}_i = \bar{\mathbf{A}}\mathbf{B}\mathbf{v}_i - \bar{\mathbf{A}}\mathbf{B}\mathbf{A}\mathbf{v}_i = \bar{\mathbf{A}}\mathbf{v}_i - \bar{\mathbf{A}}(\lambda_i \mathbf{v}_i) = (1-\lambda_i)\,\mathbf{w}_i. \tag{6}$$

Thus $\mathbf{w}_i$ is an eigenvector of $\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}$, corresponding to the eigenvalue $(1-\lambda_i)$.

Also, if $\mathbf{v}_0 \in \mathscr{V}_0$, since $\mathscr{R}(\bar{\mathbf{A}}) = \mathscr{K}(\mathbf{A}) \supset \mathscr{V}_0$, we have

$$\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}\mathbf{v}_0 = \bar{\mathbf{A}}\mathbf{B}\mathbf{v}_0 = \bar{\mathbf{A}}\mathbf{v}_0 = \mathbf{v}_0. \tag{7}$$

Hence $\mathbf{v}_0$ is also an eigenvector of $\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}$, corresponding to the eigenvalue 1. Clearly, therefore, the eigenspaces of $\bar{\mathbf{A}}\mathbf{B}\bar{\mathbf{A}}$ are $\mathscr{V}_0$ and $\mathscr{W}_i$ $(i = 1, \ldots, r)$, excluding the case $\lambda_i = 1$, by virtue of the decomposition (5).

REFERENCES

HOTELLING, H. (1936). Relations between two sets of variates. *Biometrika* **28**, 321–77.

JAMES, A. T. (1957). The relationship algebra of an experimental design. *Ann. Math. Statist.* **28**, 993–1002.

MANN, H. B. (1960). The algebra of a linear hypothesis. *Ann. Math. Statist.* **31**, 1–15.

WILKINSON, G. N. (1970). A general recursive procedure for analysis of variance. *Biometrika* **57**, 19–46.

YATES, F. (1936). Incomplete randomised blocks. *Ann. Eugen.* **7**, 121–40.

[*Received October* 1970. *Revised January* 1971]

Some key words: Geometrical interpretation of nonorthogonal analysis of variance; Canonical efficiency factor; Canonical components of regression; Relationship algebra of experimental design; Lattice design; Balance in experimental design.