

GENERALIZED DISPERSION MATRICES FOR COVARIANCE
STRUCTURAL ANALYSIS*

Harold V. Henderson¹ and Shayle R. Searle²

ABSTRACT

Methods for deriving eigenvalues and the inverse of a dispersion matrix for balanced-data variance-components models are extended to linear combinations of commutative matrices and to linear combinations of powers of a matrix or Kronecker products of powers of matrices. The result for the inverse exploits relationships with its spectral decomposition and so also requires that the matrices be simple. Applications in covariance structural analysis for (real symmetric) structured dispersion matrices and for other situations are discussed.

¹Biometrics Section, Ruakura Agricultural Research Centre,
Private Bag, Hamilton, New Zealand

²Biometrics Unit, Cornell University, Ithaca NY 14853, USA

*Paper No. BU-684-M in the Biometrics Unit, November 1984

1. INTRODUCTION

Covariance structural analysis is a generic term used when a population dispersion (variance-covariance) matrix, \tilde{V} , is assumed to have an a priori pattern (in addition to symmetry) and structure. One particular, but nevertheless quite general, structure occurring in such diverse situations as variance component estimation, experimental design and psychometrics, is considered, for example, by Anderson [1969, 1970] and Mukherjee [1976]. It is where \tilde{V} is taken as a linear combination of a set, $\{K_i\}$, of linearly independent matrices which are not functions of the parameters $\{\theta_i\}$:

$$\tilde{V} = \sum_{i=1}^c \theta_i K_i \quad (1)$$

(By linearly independent we mean that $\sum_{i=1}^c \theta_i K_i = 0$ implies that $\theta_i = 0$ for all $i=1, \dots, c$.) For general \tilde{V} of order $n \times n$ there are at most n^2 such K_i 's, so that $c \leq n^2$, and for \tilde{V} being a dispersion matrix and hence symmetric, $c \leq \frac{1}{2}n(n+1)$.

Of special interest is the case when $\tilde{V} = \sum_{i=1}^c \theta_i K_i$ and $\tilde{V}^{-1} = \sum_{i=1}^c \tau_i K_i$ have the same structure, i.e., are linear combinations (possibly with some zero weights) of the same set of K_i 's. This is discussed by Searle and Henderson [1979] for variance component models, as is now indicated.

1.1. Variance component models

Dispersion matrices for variance component models are special

cases of (1). When data from such models have equal numbers of observations in the subclasses (i.e., balanced-data), each K_i of (1) can be expressed as a Kronecker product of \underline{I} - and \underline{J} -matrices where in general \underline{I}_a is an identity matrix of order a, and \underline{J}_n is a square matrix of order n with every element unity. For example, consider the model for the two-way crossed classification with interaction:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk} \quad (2)$$

with $i=1, \dots, a$, $j=1, \dots, b$ and $k=1, \dots, n$. When all terms, except μ , are random, with zero means, all covariances zero, and with variances denoted by σ^2 with subscript corresponding to the term concerned, the dispersion matrix of y is

$$\begin{aligned} \underline{V} = & \sigma_e^2 (\underline{I}_a \otimes \underline{I}_b \otimes \underline{I}_n) + \sigma_\alpha^2 (\underline{I}_a \otimes \underline{J}_b \otimes \underline{J}_n) + \sigma_\beta^2 (\underline{J}_a \otimes \underline{I}_b \otimes \underline{J}_n) \\ & + \sigma_{\alpha\beta}^2 (\underline{I}_a \otimes \underline{I}_b \otimes \underline{J}_n) ; \end{aligned} \quad (3)$$

where \otimes denotes the operation of Kronecker product. Although $\underline{I}_a \otimes \underline{I}_n = \underline{I}_{an}$ and $\underline{J}_a \otimes \underline{J}_n = \underline{J}_{an}$, it is more convenient to work with (3) as shown. In addition to having the terms in \underline{V} of (3) with different weights, \underline{V}^{-1} has a term involving $(\underline{J}_a \otimes \underline{J}_b \otimes \underline{J}_n)$.

Searle and Henderson [1979] develop a simple methodology for deriving the eigenvalues, determinant and inverse of (3) and its generalizations to balanced p-way crossed and/or nested classifications. Similar ideas have also been discussed independently by Nelder [1965a,b] and Smith and Hocking [1978]. In the p-way case, with $\underline{J}^0 = \underline{I}$, each K_i of (1) is a Kronecker

product of powers (0 and 1) of p J -matrices, not necessarily of the same order. Then (1) is

$$V_{\sim p} = \sum_{\substack{\tilde{i} \\ i=0}}^1 \theta_{\tilde{i}} \left(J_{\sim p}^{i_p} \otimes J_{\sim p-1}^{i_{p-1}} \otimes \dots \otimes J_{\sim 1}^{i_1} \right) \quad (4)$$

where some of the θ 's are σ^2 's and the others are zero, e.g., in (3) four θ 's are σ^2 's and the other four are zero. The exponents in (4), i_p, \dots, i_1 , are also used as subscripts to θ where, for notational convenience and readability only, they are represented as a vector \tilde{i} . These subscripts are in reverse natural order to facilitate using the complete subscript to each θ as a binary number (since every index i_r for $r=p, p-1, \dots, 1$ takes only the values 0 or 1), and the summation $\sum_{\substack{\tilde{i} \\ i=0}}^1$ represents the multiple summation $\sum_{i_p=0}^1 \dots \sum_{i_1=0}^1$ and so notationally is summation over the first 2^p nonnegative binary numbers from 00...0 to 11...1.

1.2. An extension

Seely [1971] considers (1) and the special case when V is a weighted sum of powers of a symmetric matrix A which is independent of $\{\theta_i\}$:

$$V_{\sim} = \sum_{i=0}^d \theta_{\tilde{i}} A^{\tilde{i}} \quad (5)$$

Gilbert [1962] expresses a circulant matrix in this form, where A is a "one-element type" (nonsymmetric) circulant. A generalization of this, and of (4), is when each J_r in (4) is replaced by A_r of

order $n_r \times n_r$, which is independent of $\{\theta_i\}$, and the summation and powers are to $\underline{d} = (d_p \cdots d_1)'$; namely

$$\underline{V}_{\underline{p}} = \sum_{\underline{i}=0}^{\underline{d}} \theta_{\underline{i}} \underline{K}_{\underline{i}} \quad \text{where} \quad \underline{K}_{\underline{i}} = \underline{A}_{\underline{p}}^{i_p} \otimes \cdots \otimes \underline{A}_{\underline{1}}^{i_1}, \quad (6)$$

and with $\underline{V}_{\underline{p}}$ and $\underline{K}_{\underline{i}}$ having order $N_{\underline{p}} = \prod_{r=1}^p n_r$. When the d_r are all equal to d , say, the subscript of \underline{i} will range through the first $(d+1)^p$ numbers of base $b = d+1$.

That $\underline{V}_{\underline{p}}$ be a dispersion matrix, with consequent symmetry and positive definiteness, is not necessary in the subsequent development. Note that under symmetry and when $p=1$, (6) reduces to (5), the case discussed by Seely [1971] in terms of a quadratic subspace of the vector space of real symmetric matrices.

In the sequel, procedures developed for deriving eigenvalues and the inverse of (4) are extended to $\underline{V} = \sum_{i=1}^c \theta_i \underline{K}_i$ in (1) when the \underline{K}_i 's commute in pairs, and then to the special case (6) and examples thereof, with applications for symmetric \underline{V} in covariance structural analysis and for nonsymmetric \underline{V} in other situations. This paper is based on chapter 6 of the thesis by Henderson [1979].

2. EIGENVALUES AND THE INVERSE OF \underline{V}

Eigenvalues have many uses in multivariate statistics; for example, they provide invariant tests of the general linear hypothesis and are expected mean squares in variance component models.

An important occurrence of the inverse of \underline{V} is as the matrix of the quadratic form in the normal probability density function.

It therefore arises in maximum likelihood estimation under normality and hence in any subsequent hypothesis testing.

The set of all eigenvalues of a square matrix is known as its spectrum. A relationship between \underline{V} , \underline{V}^{-1} and its spectrum is exhibited in the spectral decomposition theorem.

Spectral decomposition theorem (Lancaster [1969, p. 63]):

Let \underline{V} be a simple matrix of order n with s distinct nonzero real or complex eigenvalues, λ_i for $i = 1, \dots, s$. Then

$$\underline{V} = \sum_{i=1}^s \lambda_i \underline{M}_i \quad \text{and} \quad \underline{V}^{-1} = \sum_{i=1}^s \lambda_i^{-1} \underline{M}_i \quad (7)$$

for

$$\sum_{i=1}^s \underline{M}_i = \underline{I}, \quad \underline{M}_i \underline{M}_j = \underline{0} \quad \text{for } i \neq j \quad \text{and} \quad \underline{M}_i^2 = \underline{M}_i \quad (8)$$

A special case of this theorem is usually given (e.g., Mukherjee [1976, p. 135] and Searle and Henderson [1979]) for dispersion matrices with \underline{V} , and hence the \underline{M}_i 's, being real and symmetric. But the theorem holds for \underline{V} being simple (diagonalizable or diagonalizable, i.e., similar to a diagonal matrix, e.g., real symmetric, hermitian, normal and idempotent matrices are simple). When \underline{V} is not simple we say \underline{V} is defective or not diagonalizable. In this case \underline{U}^{-1} does not exist in $\underline{V}\underline{U} = \underline{U}\underline{D}$, where \underline{D} is the diagonal matrix of eigenvalues of \underline{V} .

We now exploit the special structure of $\underline{V} = \sum_{i=1}^c \theta_i \underline{K}_i$ and its relationship to the spectral decomposition.

2.1. The main result

The eigenvalues and the inverse of $\tilde{V} = \sum_{i=1}^c \theta_i K_i$ in (1), when certain conditions apply to the K_i 's, are given by the following two-part algorithm, which is derived in succeeding paragraphs.

Algorithm:

Let $\tilde{V} = \sum_{i=1}^c \theta_i K_i$, of order $n \times n$, where K_i has m_i distinct eigenvalues with $m = \max(m_1, \dots, m_c)$.

Eigenvalues of \tilde{V} : Suppose the K_i 's are simultaneously Jordanable (commute in pairs). Then the vector of m (possibly) distinct eigenvalues of \tilde{V} is

$$\tilde{\lambda} = T\theta \quad , \quad (9)$$

where the rows of T are the m distinct rows of $L = \{l_{ij}\}$ with l_{ij} being the i th eigenvalue of K_j , and $\theta' = [\theta_1 \dots \theta_c]$.

Inverse of \tilde{V} : When

- (i) \tilde{V} has no zero eigenvalues and so is nonsingular,
- (ii) the K_i 's are simultaneously diagonalizable (commute in pairs and are simple),
- (iii) \tilde{V} is in the quadratic subspace of simple matrices generated by the linearly independent K_i 's,
- (iv) $m = c$, ensuring that T is square and nonsingular and
- (v) μ is defined as the vector of reciprocals of $\tilde{\lambda} = T\theta$ in (9),

then the inverse of \tilde{V} is:

$$\underline{V}^{-1} = \sum_{i=1}^c \tau_i \underline{K}_i \quad \text{where } \underline{\tau} = \underline{T}^{-1} \underline{\mu} \quad (10)$$

2.2. Eigenvalues of \underline{V} for simultaneously Jordanable \underline{K}_i 's

We are interested in situations where the eigenvalues of $\underline{V} = \sum_{i=1}^c \theta_i \underline{K}_i$ are linear combinations of those of the \underline{K}_i 's. Such is the case when the \underline{K}_i 's are simultaneously Jordanable. That is, for $j=1, \dots, c$,

$$\underline{K}_j = \underline{P} \underline{J}_j \underline{Q}' \quad \text{where } \underline{Q}' \underline{P} = \underline{I} \quad (11)$$

and \underline{J}_j is a Jordan normal form of \underline{K}_j - an "almost diagonal" matrix with the eigenvalues $\{l_{ij}\}$ of \underline{K}_j on the diagonal, ones or zeros on the superdiagonal and zeros elsewhere. (When \underline{K}_j is also simple \underline{J}_j is diagonal). Then \underline{V} is also Jordanized by the same matrices \underline{P} and \underline{Q} :

$$\underline{V} = \underline{P} \left(\sum_{j=1}^c \theta_j \underline{J}_j \right) \underline{Q}' \quad (12)$$

The i th eigenvalue of \underline{V} , from (12), is the linear combination

$$l_i = \sum_{j=1}^c l_{ij} \theta_j = [l_{i1} \dots l_{ic}] \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_c \end{bmatrix} \equiv \underline{l}_i' \underline{\theta} ; \quad (13)$$

hence the vector of the n eigenvalues, to be denoted by $\underline{\omega}$, is

$$\underline{\omega} = \begin{bmatrix} l_1' \\ \vdots \\ l_n' \end{bmatrix} \underline{\theta} \equiv \underline{L} \underline{\theta} \quad (14)$$

Reductions in the order of $\underline{\omega}$ and \underline{L} may be accomplished by

considering only distinct eigenvalues as presented in (9), the statement of the algorithm, which follows immediately from (14).

The relationship between the number of distinct eigenvalues of \tilde{V} and nonzero parameters, θ_i (zero θ 's and corresponding columns of \tilde{T} can be deleted), has important statistical consequences. When they are equal \tilde{T} is square and nonsingular and by the criteria of Anderson [1969] and Miller [1973], explicit maximum likelihood solutions for the parameters exist. This work is further extended and applied by Miller [1977], Szatrowski [1978, 1980], Szatrowski and Miller [1980] and Rubin and Szatrowski [1982].

2.3. Inverse of \tilde{V} for simultaneously diagonalizable K_i 's

The inverse of $\tilde{V} = \sum_{i=1}^c \theta_i K_i$ is, of course, still available from its spectral decomposition (7). But it is often instructive, see, for example, Anderson [1969, 1970], to express \tilde{V}^{-1} directly as a weighted sum of the K_i 's:

$$\tilde{V}^{-1} = \sum_{i=1}^c \tau_i K_i \quad (15)$$

When this is so we have

$$\begin{aligned} \tilde{I} = \tilde{V}\tilde{V}^{-1} &= \left(\sum_{i=1}^c \theta_i K_i \right) \left(\sum_{i=1}^c \tau_i K_i \right) \\ &= \sum_{i=1}^c \theta_i \tau_i K_i^2 + \sum_{i \neq j} \theta_i \tau_j K_i K_j \end{aligned} \quad (16)$$

One way of ensuring this is for \tilde{V} to be in a system that is closed under powers and products of K_i 's or \tilde{V} 's of the same form (viewing \tilde{V}^{-1} in (15) as another \tilde{V}). This is formalized in the following notion of a quadratic subspace.

Quadratic subspaces

A quadratic subspace of the vector space of real symmetric matrices was introduced by Seely [1971, p. 711], and subsequently called a Jordan algebra by Seely [1977]. We extend this to a vector space of structured or patterned matrices denoted by \mathcal{Q} , e.g., symmetric or circulant matrices.

Definition A subspace \mathcal{B} of \mathcal{Q} with the property that $V \in \mathcal{B}$ implies $V^2 \in \mathcal{B}$ is said to be a quadratic subspace of \mathcal{Q} . (17)

The properties developed by Seely [1971] carry over to this more general situation (deleting references to symmetry). For example, if $V \in \mathcal{B}$, then (a) $V^k \in \mathcal{B}$ for each positive integer k , (b) so is its Moore-Penrose inverse V^+ and hence when V is nonsingular so is V^{-1} and (c) there exists a basis for \mathcal{B} (which generates \mathcal{B}) consisting of idempotent matrices.

Further, when the elements of \mathcal{B} commute, the subspace \mathcal{B} is called a commutative quadratic subspace and then by lemma 6 of Seely [1971] the idempotent M_i 's of (7) and (8) constitute the unique idempotent basis, apart from indexing.

It is sufficient that V be in the quadratic subspace generated by the idempotent M_i 's or all the K_i 's for $V = \sum_{i=1}^c \theta_i K_i$ to imply that $V^{-1} = \sum_{i=1}^c \tau_i K_i$ for some values of τ_i . The spectral decomposition of V in (7) and (8) is a special case of this. In general, this requirement simply ensures that we have all the K_i 's, even those with zero coefficients in V . In contrast to Anderson

[1969] we describe \tilde{V} and \tilde{V}^{-1} as having the same structure even when some of the coefficients in either (or both) of \tilde{V} and \tilde{V}^{-1} are zero. By allowing this generalization we introduce the subtlety that when θ_i is zero, τ_i need not be zero, so that although \tilde{V} and \tilde{V}^{-1} are both linear combinations of all the K_i 's, the particular subsets with nonzero weights need not be the same. This is so with (3), as detailed by Searle and Henderson [1979]. Even though the K_i 's with nonzero θ_i 's in \tilde{V} may not be all the same as those with nonzero τ_i 's in \tilde{V}^{-1} we still describe \tilde{V} and \tilde{V}^{-1} as having the same structure.

When \tilde{V} and \tilde{V}^{-1} have the same form as just defined the algorithm for \tilde{V}^{-1} given in (10) is established by exploiting the relation between the spectral decompositions of \tilde{V} and \tilde{V}^{-1} ; hence the requirement that the K_i 's also be simple. Specifically, define $\underline{\mu}$ as the vector of reciprocal elements of $\underline{\lambda} = T\underline{\theta}$, i.e. $\mu_i = \lambda_i^{-1}$. Then because $\tilde{V} = \sum_{i=1}^c \theta_i K_i$ and $\tilde{V}^{-1} = \sum_{i=1}^c \tau_i K_i$ have the same structure $\underline{\mu} = T\underline{\tau}$. The K_i 's and hence the columns of T are linearly independent and so $\underline{\tau} = T^{-1}\underline{\mu}$.

Having established a general algorithm for $\tilde{V} = \sum_{i=1}^c \theta_i K_i$, we now consider the special case (6) and examples thereof.

3. EIGENVALUES AND THE INVERSE OF \tilde{V}_P

The special case (6) covers a wide range of application.

Here the K_i 's commute, have special structure and are all evident;

by virtue of their being Kronecker products of powers of matrices. The eigenvalues and the inverse of $V_{\sim p}$ are given by the following algorithm, which is derived from the general algorithm, (9) and (10), in the next two sections. This result generalizes the algorithm of Searle and Henderson [1979] for (4), the result for the eigenvalues of (5) in Seely [1971, pp. 711-712] and that given by Gilbert [1962] for circulants.

This result reduces finding the $B_p = \prod_{r=1}^p b_r$ (possibly) distinct eigenvalues of $V_{\sim p}$ of order $N_p = \prod_{r=1}^p n_r$ to performing matrix and vector multiplication of order B_p and evaluating the b_r distinct eigenvalues of the p $A_{\sim r}$ -matrices of order n_r . This can represent considerable computational reduction compared to the alternative of directly decomposing $V_{\sim p}$. In fact, a recursive algorithm is available, as given in Section 3.3.

Similarly, inverting $V_{\sim p}$, of order N_p is reduced to matrix and vector multiplication of order B_p and the inversion of p Vandermonde matrices of order b_r .

Algorithm: We consider $V_{\sim p}$ as detailed in (6),

$$V_{\sim p} = \sum_{i=0}^d \theta_{\sim i} K_{\sim i} \quad \text{with} \quad K_{\sim i} = A_{\sim p}^i \otimes \dots \otimes A_{\sim 1}^i \quad (18)$$

of order $N_p = \prod_{r=1}^p n_r$, where $A_{\sim r}$ is square of order n_r

and has b_r distinct eigenvalues, $a_{1r}, \dots, a_{b_r r}$.

Eigenvalues: Then the vector of the $B_p = \prod_{r=1}^p b_r$ (possibly) distinct

eigenvalues of $V_{\sim p}$ is given by $\lambda_{\sim p} = T_{\sim p} \theta_{\sim p}$, where

$$\lambda_{\sim p} \equiv \begin{bmatrix} \lambda_{00} \dots 0 \\ \lambda_{00} \dots 1 \\ \vdots \\ \lambda_{d_r} \dots d_1 \end{bmatrix}, \quad \theta_{\sim p} = \begin{bmatrix} \theta_{00} \dots 0 \\ \theta_{00} \dots 1 \\ \vdots \\ \theta_{d_r} \dots d_1 \end{bmatrix}, \quad T_{\sim p} = C_{\sim p} \otimes \dots \otimes C_1 \quad (19)$$

and

$$\begin{aligned} C_r &\equiv C_r(a_{1r}, \dots, a_{b_r r}) \\ &= \{a_{ir}^{j-1}\} \quad \text{for } i=1, \dots, b_r \text{ and } j=1, \dots, d_r+1 \\ &= \begin{bmatrix} 1 & a_{1r} & \dots & a_{1r}^{d_r} \\ \vdots & \vdots & & \vdots \\ 1 & a_{b_r r} & \dots & a_{b_r r}^{d_r} \end{bmatrix} \end{aligned} \quad (20)$$

is a Vandermonde matrix.

Inverse: When

- (i) $V_{\sim p}$ has no zero eigenvalues,
- (ii) each $A_{\sim r}$ (and hence $K_{\sim i}$) is simple,
- (iii) $V_{\sim p}$ is in the quadratic subspace generated by the $K_{\sim i}$'s,

and

- (iv) $b_r = d_r + 1$ (so that $C_{\sim r}$ and hence T is nonsingular),

then on defining

$$\mu_{\sim p} \text{ as the vector of reciprocals of elements of } \lambda_{\sim p}, \quad (21)$$

the coefficients in the inverse of $V_{\sim p}$,

$$V_{\sim p}^{-1} = \sum_{i=0}^d \tau_{\sim i} K_{\sim i} \quad \text{with } K_{\sim i} = A_{\sim p}^{i_p} \otimes \cdots \otimes A_{\sim 1}^{i_1} \quad (22)$$

are given by

$$\tau_{\sim p} = T_{\sim p}^{-1} \mu_{\sim p} \quad (23)$$

using $T_{\sim p}^{-1} = C_{\sim p}^{-1} \otimes \cdots \otimes C_{\sim 1}^{-1}$.

3.1. Eigenvalues of $V_{\sim p}$

By definition (18), $V_{\sim p}$ is a weighted sum of Kronecker products of p matrices. When all the d_r are equal Lancaster [1969, p. 259] for $p=2$ and Searle and Henderson [1979] for any $p \geq 1$ have theorems for determining the eigenvalues of such a sum. We now extend this to the general case of unequal d_r . The extension is straightforward and is stated without proof. The result relies on the Jordan normal form of $A_{\sim r}$, as does (13), of which this is a special case.

Theorem: Suppose $A_{\sim r}$ is square of order n_r with eigenvalues $l_{1r}, \dots, l_{n_r r}$. Then the N_p eigenvalues of $V_{\sim p}$ of (18) are polynomials of the l 's with the θ 's as coefficients:

$$\lambda_{\sim t} = \sum_{i=0}^d \theta_{\sim i} \left(\prod_{r=1}^p l_{t_r}^{i_r} \right), \quad (24)$$

where t_p, \dots, t_1 is denoted by $\sim t$ and where each l_{t_r} takes, in turn, the values of $l_{1r}, \dots, l_{n_r r}$.

We now express (24) in matrix notation, akin to (13). Observe that $\lambda_t = (\underline{l}_p \otimes \dots \otimes \underline{l}_p)' \theta_{\underline{p}}$ where $\underline{l}'_r = \begin{bmatrix} 1 & l_{t_r} & \dots & l_{t_r}^{d_r} \end{bmatrix}$ and so the N_p eigenvalues are given by

$$\omega_{\underline{p}} = (\underline{L}_p \otimes \dots \otimes \underline{L}_1) \theta_{\underline{p}} \quad \text{where} \quad \underline{L}_r = \begin{bmatrix} 1 & l_{1r} & \dots & l_{1r}^{d_r} \\ 1 & l_{2r} & \dots & l_{2r}^{d_r} \\ \vdots & \vdots & & \vdots \\ 1 & l_{n_r r} & \dots & l_{n_r r}^{d_r} \end{bmatrix} \quad (25)$$

is of order $n_r \times (d_r + 1)$. The statement in the algorithm for distinct eigenvalues follows directly from (25) in the same way (9) does from (14), where \underline{C}_r has the distinct rows of \underline{L}_r .

3.2. Inverse of $V_{\underline{p}}$

The algorithm for $V_{\underline{p}}^{-1}$ follows as a special case of that for $V_{\underline{p}}^{-1}$ of (10); the subscript p being added to all terms. The special structure of $\underline{T}_{\underline{p}} = \underline{C}_{\underline{p}} \otimes \dots \otimes \underline{C}_{\underline{1}}$ ensures that it is nonsingular. This follows because the Vandermonde matrix \underline{C}_r is nonsingular; its determinant with distinct a_{ir} 's being $|\underline{C}_r| = \prod_{i>j} (a_{ir} - a_{jr})$ and hence nonzero.

3.3. A recursive property

A recursive property of the eigenvalues of $\lambda_{\underline{p}} = \underline{T}_{\underline{p}} \theta_{\underline{p}}$ is worthy of note. From (19)

$$\underline{T}_{\underline{p}} = \underline{C}_{\underline{p}} \otimes \underline{T}_{\underline{p}-1} = \left\{ a_{ip}^{j-1} \underline{T}_{\underline{p}-1} \right\} = \begin{bmatrix} \underline{T}_{\underline{p}-1} & a_{1p} \underline{T}_{\underline{p}-1} & \dots & a_{1p}^{d_p} \underline{T}_{\underline{p}-1} \\ \vdots & \vdots & & \vdots \\ \underline{T}_{\underline{p}-1} & a_{b,p} \underline{T}_{\underline{p}-1} & \dots & a_{b,p}^{d_p} \underline{T}_{\underline{p}-1} \end{bmatrix} \quad (26)$$

for $i=1, \dots, b_p$ and $j=1, \dots, d_p+1$.

Define $\theta_{i,p-1} = \{\theta_{i_1, p-1} \dots i_1\}$ for $i=0, \dots, d_p$ with $i_r=0, \dots, d_r$ for $r=1, \dots, p-1$, observing that

$$\theta_{\sim p} = \begin{bmatrix} \theta_{\sim 0, p-1} \\ \vdots \\ \theta_{\sim d_p, p-1} \end{bmatrix}, \quad (27)$$

and similarly for $\lambda_{\sim p}$. The vector of (possibly) distinct eigenvalues of $V_{\sim p}$ is given by

$$\lambda_{\sim p} = T_{\sim p} \theta_{\sim p} \text{ with its } i\text{th subvector} \quad (28)$$

$$\lambda_{i,p-1} = T_{\sim p} \theta_{\cdot, p-1, i}$$

on defining

$$\theta_{\cdot, p-1, i} = \sum_{r=0}^{d_p} a_{ip}^r \theta_{r, p-1} \quad (29)$$

as the i th weighted sum of $\theta_{r, p-1}$'s.

For example, Searle and Henderson [1979] give (26) for the special case (4), with examples, where

$$T_{\sim p} = \begin{bmatrix} T_{\sim p-1} & 0 \\ T_{\sim p-1} & n_{p-1} T_{\sim p-1} \end{bmatrix}, \quad \theta_{\sim p} = \begin{bmatrix} \theta_{\sim 0, p-1} \\ \theta_{\sim 1, p-1} \end{bmatrix}. \quad (30)$$

In this case (28) gives the eigenvalues of $V_{\sim p}$ as

$$\lambda_{\sim p} = \begin{bmatrix} \lambda_{\sim 0, p-1} \\ \lambda_{\sim 1, p-1} \end{bmatrix} = \begin{bmatrix} T_{\sim p-1} \theta_{\cdot, p-1, 0} \\ T_{\sim p-1} \theta_{\cdot, p-1, 1} \end{bmatrix}, \quad (31)$$

4. SPECIFIC APPLICATIONS

We now illustrate the algorithm for specific examples of $V_{\sim p}$ in (18), of interest in covariance structural analysis.

4.1 Variance component models

Several examples of $V_{\sim p}$ being the dispersion matrix for a variance component model for balanced data, similar to (3) and its generalization (4), are presented by Searle and Henderson [1979], using a special case of the algorithm specific to variance component models. Nelder [1965a,b] also presents a number of examples in a detailed discussion of the analysis of randomized experiments with orthogonal block structure. Extensions to this work are developed in Speed [1981, 1983], Houtman and Speed [1983] and references therein. Smith and Hocking [1978] also independently develop similar methodology for variance component models for balanced data. Searle and Rudan [1973], Wansbeek [1982] and Bunney and Kissling [1984] discuss inverting a dispersion matrix for unbalanced data.

4.2. Quasi-circumplex correlation matrix

Consider the pattern matrix

$$V_{\sim 2} = \begin{bmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{bmatrix} \quad (32)$$

$$\begin{aligned} &= a(I_{\sim 2} \otimes I_{\sim 2}) + b(I_{\sim 2} \otimes E_{\sim 2}) + c(E_{\sim 2} \otimes I_{\sim 2}) + d(E_{\sim 2} \otimes E_{\sim 2}) \\ &= \sum_{i,j=0}^1 \theta_{ij} (E_{\sim 2}^i \otimes E_{\sim 2}^j) \end{aligned} \quad (33)$$

where $E_{\sim 2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\theta' = [a \ b \ c \ d]$.

The eigenvalues of \underline{E}_2 are ± 1 , so using (19) with $\underline{C}_2 = \underline{C}_1 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, those of \underline{V}_2 are

$$\begin{aligned} \underline{\lambda}_2 &= \underline{T}_2 \underline{\theta}_2 = (\underline{C}_2 \otimes \underline{C}_1) \underline{\theta}_2 \\ &= \begin{bmatrix} \lambda_{00} \\ \lambda_{01} \\ \lambda_{10} \\ \lambda_{11} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a - b - c + d \\ a + b - c - d \\ a - b + c - d \\ a + b + c + d \end{bmatrix} \quad (34) \end{aligned}$$

To express $\underline{V}^{-1} = \sum_{i,j=0}^1 \tau_{ij} (\underline{E}_2^i \otimes \underline{E}_2^j)$, we use $\underline{\tau}_2 = \underline{T}_2^{-1} \underline{\mu}_2$, where $\underline{\mu}_2$ is the vector of reciprocal elements of $\underline{\lambda}_2$. Since $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$,

$$\begin{aligned} \underline{\tau}_2 &= \underline{T}_2^{-1} \underline{\mu}_2 = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \underline{\mu}_2 \\ &= \frac{1}{4} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \mu_{00} \\ \mu_{01} \\ \mu_{10} \\ \mu_{11} \end{bmatrix} \quad (35) \end{aligned}$$

When $\underline{\theta}'_2 = [1 \ \rho_1 \ \rho_2 \ \rho_1]$, \underline{V}_2 is a 4×4 Guttman [1954] quasi-circumplex correlation matrix with eigenvalues $\underline{\lambda}'_2 = [1 - \rho_2 \ 1 - \rho_2 \ 1 + \rho_2 - 2\rho_1 \ 1 + 2\rho_1 + \rho_2]$. The consequent spectral decomposition is given by Mukherjee [1976, p. 139] although he does not recognize the Kronecker product structure.

In the previous examples the \underline{K}_i 's have enjoyed both special power and Kronecker product structure. The next two applications have only power structure.

4.3. A spectral matrix

Consider the dispersion matrix (with $a > b$) discussed by Mukherjee [1976, p. 136],

$$\tilde{V} = \begin{bmatrix} a+b & 0 & a-b \\ 0 & 2a & 0 \\ a-b & 0 & a+b \end{bmatrix}, \quad (36)$$

which can be decomposed by inspection into its spectral decomposition:

$$\tilde{V} = 2a \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + 2b \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \equiv \lambda_1 \underline{M}_1 + \lambda_2 \underline{M}_2 \quad (37)$$

where $\lambda_1 = 2a$ has multiplicity 2. Its inverse has then, of course, the same form:

$$\tilde{V}^{-1} = \lambda_1^{-1} \underline{M}_1 + \lambda_2^{-1} \underline{M}_2 = \frac{1}{4ab} \begin{bmatrix} a+b & 0 & b-a \\ 0 & 2b & 0 \\ b-a & 0 & a+b \end{bmatrix}. \quad (38)$$

To illustrate the algorithm consider the more obvious decomposition:

$$\tilde{V} = (a+b)\underline{I} + (a-b)\underline{K} \quad \text{where} \quad \underline{K} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \quad (39)$$

Then \tilde{V} is in a quadratic subspace generated by \underline{I} and \underline{K} and the algorithm to express \tilde{V}^{-1} in terms of $\underline{I} = \underline{K}^0$ and \underline{K} is applicable. The eigenvalues of \underline{K} are 1, 1 and -1, so (9) gives

$$\tilde{\lambda} = \tilde{T}\tilde{\theta} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a+b \\ a-b \end{bmatrix} = \begin{bmatrix} 2a \\ 2b \end{bmatrix} \quad (40)$$

as in (37). The coefficients in \tilde{V}^{-1} are

$$\tilde{\tau} = \tilde{T}^{-1}\tilde{\mu} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1/2a \\ 1/2b \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1/a + 1/b \\ 1/a - 1/b \end{bmatrix} = \frac{1}{4ab} \begin{bmatrix} a+b \\ b-a \end{bmatrix} \quad (41)$$

Hence, \tilde{V}^{-1} is

$$\tilde{V}^{-1} = \frac{a+b}{4ab} \tilde{I} + \frac{b-a}{4ab} \tilde{K} \quad ,$$

as in (38).

4.4. Circulants

A circulant is a square matrix having the form

$$\tilde{C} = \tilde{C}(c_0, \dots, c_{n-1}) = \begin{bmatrix} c_0 & c_1 & \dots & c_{n-1} \\ c_{n-1} & c_0 & \dots & c_{n-2} \\ \vdots & \vdots & & \vdots \\ c_1 & c_2 & \dots & c_0 \end{bmatrix} \quad (42)$$

The inverse of a nonsingular circulant is itself a circulant. A number of methods for obtaining the inverse are available; see Davis [1979] and Searle [1979], who also gives a new method involving recurrence equations, and Feinsilver [1984].

An earlier method, from Gilbert [1962], of transforming to diagonal form, inverting, and then transforming back to original form is a special case of our general algorithm applied to circulants.

This is achieved by defining the "one-element type" circulant $\underline{A} = \underline{C}(0, 1, \dots, 0)$ and expressing \underline{C} in the form of (5) as

$$\underline{C} = \sum_{i=0}^{n-1} c_i \underline{A}^i . \quad (43)$$

The eigenvalues of \underline{A} are well-known and the application of the algorithm is straightforward.

Symmetric circulants have been considered in statistical applications, e.g., as the dispersion matrix for certain cyclic partially balanced designs, by Wise [1955], Srivastava [1966], Olkin and Press [1969], T. W. Anderson [1969] and D. A. Anderson [1972]. T. W. Anderson [1969] presents the 6×6 case as illustration, but does not appeal to the power structure afforded by (43).

REFERENCES

- Anderson, D. A. [1972]. Designs with partial factorial balance. Ann. Math. Statist. 43, 1333-1341.
- Anderson, T. W. [1969]. Statistical inference for covariance matrices with linear structure. Multivariate Analysis-II. (P.R. Krishnaiah, Ed.), Academic Press, New York, 55-66.
- Anderson, T. W. [1970]. Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices. Essays in Probability and Statistics, University of North Carolina Press, Chapel Hill, 1-24.
- Bellman, Richard [1970]. Introduction to Matrix Analysis, second edition. McGraw-Hill, New York.
- Bonney, G. E. and Kissling, G. E. [1984]. On the inverse of a patterned covariance matrix. J. Amer. Statist. Assoc., 79, 722-723.
- Davis, Philip J. [1979]. Circulant Matrices. Wiley, New York.
- Feinsilver, Philip [1984]. Circulants, inversion of circulants, and some related matrix algebra. Linear Algebra and Appl., 56, 29-43.
- Gilbert, T. L. [1962]. Inverse overlap matrix for periodic arrays of atoms. J. Math. Phys., 3, 107-117.
- Henderson, Harold V. [1979]. Dispersion, patterned and structured matrices: vec operators and related topics. PhD Thesis, Cornell University. 182 pp.
- Houtman, A. M. and Speed, T. P. [1983]. Balance in designed experiments with orthogonal block structure. Ann. Statist., 11, 1069-1085.
- Lancaster, P. [1969]. Theory of Matrices. Academic Press, New York.

- Miller, J. J. [1973]. Asymptotic properties and computation of maximum likelihood estimates in the mixed model of the analysis of variance. Technical Report No. 12, Stanford University.
- Miller, John J. [1977]. Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance. Ann. Statist., 5, 746-762.
- Mukherjee, B. N. [1976]. Techniques of covariance structural analysis. Austral. J. Statist., 18, 131-150.
- Nelder, J. A. [1965a]. The analysis of randomized experiments with orthogonal block structure, I. Block structure and the null analysis of variance. Proc. Roy. Soc. (London) Ser. A, 273, 147-162.
- Nelder, J. A. [1965b]. The analysis of randomized experiments with orthogonal block structure, II. Treatment structure and the general analysis of variance. Proc. Roy. Soc. (London) Ser. A, 273, 163-178.
- Olkin, I. and Press, S. J. [1969]. Testing and estimation for a circular stationary model. Ann. Math. Statist., 40, 1358-1373.
- Rubin, Donald B. and Szatrowski, Ted [1982]. Finding maximum likelihood estimates of patterned covariance matrices by the EM algorithm. Biometrika, 69, 657-660.
- Searle, S. R. [1971]. Linear Models. Wiley, New York.
- Searle, S. R. [1979]. On inverting circulant matrices. Linear Algebra and Appl., 25, 77-89.
- Searle, S. R. and Henderson, H. V. [1979]. Dispersion matrices for variance component models. J. Amer. Statist. Assoc., 74, 465-470.
- Searle, S. R. and Rudan, J. W. [1973]. Wanted: an inverse matrix. Cummun. Statist., 2, 155-166.

- Seely, Justus [1971]. Quadratic subspaces and completeness. Ann. Math. Statist., 42, 710-721.
- Seely, Justus [1977]. Minimal sufficient statistics and completeness for multivariate normal families. Sankhyā, Ser. A, 39, 170-185.
- Smith, David W. and Hocking, R. R. [1978]. Maximum likelihood analysis of the mixed model: the balanced case. Commun. Statist.-Theor. Meth. A7, 1253-1266.
- Speed, T. P. [1981]. Linear models and the analysis of variance. Technical Report, Department of Mathematics, University of Western Australia. 60 pp.
- Speed, T. P. [1983]. General Balance. Encyclopedia of Statistical Sciences (Kotz, S. and Johnson, N. L., Ed.), Vol. 3, Wiley, New York. 320-326.
- Srivastava, J. N. [1966]. On testing of hypotheses regarding a class of covariance structures. Psychometrika, 31, 147-164.
- Szatrowski, Ted H. [1978]. Explicit solutions, one iteration convergence and averaging in the multivariate normal estimation problem for patterned means and covariances. Ann. Inst. Statist. Math., 30, 81-88.
- Szatrowski, Ted H. [1980]. Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. Ann. Statist., 8, 802-810.
- Szatrowski, Ted H. and Miller, John J. [1980]. Explicit maximum likelihood estimates from balanced data in the mixed model of the analysis of variance. Ann. Statist., 8, 811-819.

Wansbeek, Tom [1982]. Another approach to inverting a covariance matrix when data are unbalanced. Commun. Statist.-Simula. Computa., 11, 583-588.

Wise, J. [1955]. The autocorrelation function and the spectral density function. Biometrika, 42, 151-159.