GENERALIZED DISPERSION MATRICES FOR COVARIANCE STRUCTURAL ANALYSIS

Harold V. Henderson* and Shayle R. Searle
Biometrics Unit, Cornell University, Ithaca, NY 14853

BU-947-M

ABSTRACT

Matrices **V** that are linear combinations of linearly independent matrices **K** are considered. Methods are given for deriving eigenroots (eigenvalues) and the inverse of **V** when the **K**-matrices are simultaneously diagonalizable and closed under multiplication, including such cases as the **K**-matrices being powers of a matrix, or Kronecker products of powers of matrices. The methods are extensions of those used for dispersion matrices for balanced-data variance-components models. 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.

*This paper was partially supported by the Mathematical Sciences Institute at Cornell University, while the first author was on leave from the Statistics and Computing Section, Ruakura Agricultural Research Center, Hamilton, New Zealand as Visiting Fellow in the Biometrics Unit at Cornell University, August – October, 1987.

Technical Report number BU-947-M in the Biometrics Unit.

1. INTRODUCTION

Covariance structural analysis is a generic term used when a population dispersion (variance-covariance) matrix, $\mathbf{V}$, is assumed to have some *a priori* pattern (in addition to symmetry) and structure. One particular, but nevertheless quite general, pattern occurring in such diverse situations as variance component estimation, experimental design and psychometrics is where there are linear relationships among the variances and covariances that constitute the elements of $\mathbf{V}$. More specifically, we consider the case where $\mathbf{V}$ is taken as a linear combination of c matrices $\mathbf{K}_1, \mathbf{K}_2, \cdots, \mathbf{K}_c$:

$$\mathbf{V} = \sum_{i=1}^{c} \theta_i \mathbf{K}_i \ , \tag{1}$$

where none of the matrices $\mathbf{K}_i$ is a function of the scalars $\theta_1, \cdots, \theta_c$. Moreover, the $\mathbf{K}_i$-matrices are taken to be linearly independent, meaning that $\mathbf{V}$ can be null only if every θ_i is zero. For $\mathbf{V}$ of order $n \times n$ there are at most n^2 such $\mathbf{K}_i$-matrices, so that $c \le n^2$; and for $\mathbf{V}$ being a dispersion matrix and hence symmetric, $c \le \frac{1}{2}n(n+1)$.

We refer to $\mathbf{V}$ of (1) as having a linear covariance pattern. Such patterns are considered, for example, by Anderson (1969,1970), Mukherjee (1976,1984) and Szatrowski (1980). They are reviewed in the encyclopedia article by Szatrowski (1985) and in the text by Takeuchi, Yanai and Mukherjee (1982, Chapter 10); and are briefly mentioned in standard texts such as Anderson (1984, p.101) and Seber (1984, p.102).

An especially important case of (1) is where the inverse of $\mathbf{V}$ has the same form as $\mathbf{V}$, namely a linear combination of the same $\mathbf{K}_i$-matrices that constitute $\mathbf{V}$, i.e.,

$$\mathbf{V}^{-1} = \sum_{i=1}^{c} \tau_i \mathbf{K}_i \tag{2}$$

for some set of scalars $\tau_1,\cdots,\tau_c$. Included here is the possibility that for some values of i, we might have $\theta_i = 0$ and $\tau_i \neq 0$ or $\theta_i \neq 0$ and $\tau_i = 0$, as well as the more usual cases of both θ_i and τ_i being non-zero, or zero together. This similarity of form of **V** and $\mathbf{V}^{-1}$ yields some interesting statistical consequences. For example, Anderson (1969), Miller (1973), and Szatrowski (1980, Theorem 4) give conditions under which these forms lead to non-iterative solutions to the maximum likelihood (under normality) equations for the θs, which are then covariance parameters. This work is further extended and applied by Miller (1977), Szatrowski (1978,1980), Szatrowski and Miller (1980) and Rubin and Szatrowski (1982). In this paper we develop procedures for deriving eigenroots and the inverse of $\mathbf{V} = \Sigma\theta_i\mathbf{K}_i$ of (1) when the $\mathbf{K}_i$-matrices are simple (diagonalizable — see Section 2.1) and are pairwise commutative. Special cases are then considered, with applications to symmetric **V** in covariance structural analysis and to non-symmetric **V** in other situations. The paper is thus an extension of Henderson (1979, Chapter 6).

2. SOME MATRIX PRELIMINARIES

We briefly recall three concepts and a theorem that are well known but pertinent to our development.

2.1 Simple matrices A matrix **A** is described as simple if there exists a non-singular matrix **P** such that $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix. And that diagonal matrix does, of course, have the eigenroots of **A** on its diagonal. Thus a simple matrix is said to be diagonalizable. Examples of simple matrices are matrices with distinct eigenroots, or matrices which are real and symmetric, or Hermitian, or normal, or circulant, or idempotent.

2.2 Simultaneous diagonalizability Simple matrices that commute in multiplication are simultaneously diagonalizable (e.g., Mirsky, 1982, p. 318). Thus if $\mathbf{AB} = \mathbf{BA}$ where $\mathbf{A}$ and $\mathbf{B}$ are both simple, then there exists a non-singular $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{AP}$ and $\mathbf{P}^{-1}\mathbf{BP}$ are, for the same $\mathbf{P}$, diagonal matrices (of eigenroots of $\mathbf{A}$ and $\mathbf{B}$, respectively).

2.3 Closure under multiplication The set of matrices $\mathbf{K}_1, \mathbf{K}_2, \cdots, \mathbf{K}_i, \cdots, \mathbf{K}_c$ are said to be closed under multiplication when every product, $\mathbf{K}_i^2$ and $\mathbf{K}_i\mathbf{K}_j$ for $i \neq j$, is a scalar multiple of some $\mathbf{K}_i$ for $i = 1,\cdots,c$.

2.4 The spectral decomposition (e.g., Lancaster, 1969, p. 63). The n eigenroots of a matrix of order n×n are known as its spectrum. Denote by s the number of distinct eigenroots $\lambda_1,\lambda_2,\cdots,\lambda_s$. of a simple, non-singular matrix $\mathbf{A}$. Then the spectral decomposition is that there exist matrices $\mathbf{M}_t$ for $t = 1,\cdots,s$ such that

$$\mathbf{A} = \sum_{t=1}^{s} \lambda_t \mathbf{M}_t \qquad \text{and} \qquad \mathbf{A}^{-1} = \sum_{t=1}^{s} \lambda_t^{-1} \mathbf{M}_t \tag{3}$$

where

$$\sum_{t=1}^{s} \mathbf{M}_t = \mathbf{I}, \quad \mathbf{M}_t^2 = \mathbf{M}_t \quad \text{and} \quad \mathbf{M}_t\mathbf{M}_{t'} = 0 \quad \text{for } t \neq t' . \tag{4}$$

Special cases of the spectral decomposition of interest in statistics are for $\mathbf{A}$ being a dispersion matrix $\mathbf{V}$, which is real and symmetric, whereupon each $\mathbf{M}_i$ of (3) and (4) is also, e.g., Mukherjee (1976, p. 135) and Searle and Henderson (1979).

3. EIGENROOTS AND INVERSES

The special structure of **V** when it has the form $\mathbf{V} = \Sigma_i \theta_i \mathbf{K}_i$ of (1) is now exploited to give expressions for the eigenroots and inverse of **V** when the $\mathbf{K}_i$-matrices are simultaneously diagonalizable and the set of such matrices is closed under multiplication. The results are given in two theorems: the first gives eigenvalues of **V** and the second gives $\mathbf{V}^{-1}$

3.1 Eigenroots

Theorem 1 For $\mathbf{V} = \sum_{i=1}^{c} \theta_i \mathbf{K}_i$ with the $\mathbf{K}_i$-matrices being linearly independent and simultaneously diagonable, let $\mathbf{e}(\mathbf{K}_i)$ be the column vector of eigenroots of $\mathbf{K}_i$ sequenced in the same order as in the diagonalization $\mathbf{P}^{-1}\mathbf{K}_i\mathbf{P}$, and let $\mathbf{L}_c$ be the full column rank matrix of those c vectors:

$$\mathbf{L}_c = [\mathbf{e}(\mathbf{K}_1) \quad \mathbf{e}(\mathbf{K}_2) \quad \cdots \quad \mathbf{e}(\mathbf{K}_c)] \; . \tag{5}$$

Then, for

$$\boldsymbol{\theta}'_c = [\theta_1 \quad \theta_2 \quad \cdots \quad \theta_c] \; , \tag{6}$$

the vector of eigenroots of **V** is

$$\mathbf{e}(\mathbf{V}) = \mathbf{L}_c \boldsymbol{\theta}_c \; . \tag{7}$$

Proof: Because the $\mathbf{K}_i$-matrices are simultaneously diagonable with

$$\mathbf{P}^{-1}\mathbf{K}_i\mathbf{P} = \mathbf{D}_i, \text{ say} \; , \tag{8}$$

V is also diagonable using the same **P**:

$$\mathbf{P}^{-1}\mathbf{V}\mathbf{P} = \sum_{i=1}^{c} \theta_i \mathbf{D}_i = \mathbf{D}, \text{ say}, \tag{9}$$

where **D** is the diagonal matrix of eigenroots of **V**. Therefore

$$\mathbf{e}(\mathbf{V}) = \mathbf{e}(\mathbf{D}) = \sum_{i=1}^{c} \boldsymbol{\theta}_i \mathbf{e}(\mathbf{D}_i) = \sum_{i=1}^{c} \boldsymbol{\theta}_i \mathbf{e}(\mathbf{K}_i) = \mathbf{L}_c \boldsymbol{\theta}_c \; .$$

Q.E.D.

Note that $\mathbf{L}_c$ has full column rank because the **K**s are linearly independent. Also, in each $\mathbf{D}_i$ of (8) and in $\mathbf{D}$ of (9) the eigenroots of the $\mathbf{K}_i$-matrices and of $\mathbf{V}$ are sequenced according to the (simultaneous) diagonalization of those matrices using **P**; and this same sequence holds in the column vectors $\mathbf{e}(\mathbf{K}_i)$ and $\mathbf{e}(\mathbf{V})$.

3.2 Inverse Suppose now that the c matrices $\mathbf{K}_i$ occurring in $\mathbf{V}$ come from a (possibly) larger set of q matrices ($q \geq c$) that is closed under multiplication. Then $\mathbf{V}$ can be expressed as

$$\mathbf{V} = \sum_{i=1}^{q} \boldsymbol{\theta}_i \mathbf{K}_i \tag{10}$$

where some of the θ_i-values in (10) may be zero. Theorem 1 still applies, in the form

$$\mathbf{e}(\mathbf{V}) = \mathbf{L}_q \boldsymbol{\theta}_q \tag{11}$$

with $\mathbf{L}_q$ of q columns used in place of $\mathbf{L}_c$ of c columns, and with $\boldsymbol{\theta}_q$ having q elements, some of which may be zero. Then $\mathbf{V}^{-1}$ is given by the following theorem.

Theorem 2 Suppose $\mathbf{V} = \sum_{i=1}^{q} \boldsymbol{\theta}_i \mathbf{K}_i$ is non-singular where the $\mathbf{K}_i$-matrices are linearly independent, simple, simultaneously diagonalizable (commutative) and the set of **K**s is closed under multiplication. Let $\mathbf{L}_q$ be the matrix of **q** (linearly independent) column vectors $\mathbf{e}(\mathbf{K}_i)$ of eigenroots of $\mathbf{K}_i$ and let $\mathbf{e}(\mathbf{V}) = [\delta_1 \;\; \delta_2 \;\cdots\; \delta_n]'$ be the vector of n eigenroots of $\mathbf{V}$. (As before, all eigenroots are ordered by the simultaneous diagonalization based on **P**.) Then

$$\mathbf{V}^{-1} = \sum_{i=1}^{q} \tau_i \mathbf{K}_i \tag{12}$$

for

$$\boldsymbol{\tau} = [\tau_1 \;\; \tau_2 \;\cdots\; \tau_q]' = \mathbf{L}_q^{+} \mathbf{e}(\mathbf{V}^{-1}) \tag{13}$$

where $\mathbf{L}_q^{+} = (\mathbf{L}_q'\mathbf{L}_q)^{-1}\mathbf{L}_q'$ is the Moore-Penrose inverse of the full column rank $\mathbf{L}_q$. And

$$\mathbf{e}(\mathbf{V}^{-1}) = [1/\delta_1 \quad 1/\delta_2 \quad \cdots \quad 1/\delta_n]' \ , \tag{14}$$

where every δ is non-zero because $\mathbf{V}$ is non-singular.

Proof: The spectral decomposition theorem applied to $\mathbf{V} = \sum_{i=1}^{q} \boldsymbol{\theta}_i \mathbf{K}_i$ yields $\mathbf{V} = \sum_{j=1}^{s} \lambda_j \mathbf{M}_j$ for the s distinct eigenroots, $\lambda_1, \cdots, \lambda_s$, of $\mathbf{V}$; and Theorem 1 gives $\mathbf{e}(\mathbf{V}) = \mathbf{L}_q \boldsymbol{\theta}_q$ as in (11). Therefore, since for non-singular $\mathbf{V}$ the spectral decomposition theorem gives $\mathbf{V}^{-1} = \sum_{j=1}^{s} (1/\lambda_j)\mathbf{M}_i$, there must be scalars $\tau_1, \tau_2, \cdots, \tau_k$ such that

$$\mathbf{V}^{-1} = \sum_{i=1}^{q} \tau_i \mathbf{K}_i$$

with, from Theorem 1

$$\mathbf{e}(\mathbf{V}^{-1}) = \mathbf{L}_q \boldsymbol{\tau} \ . \tag{15}$$

But with $\mathbf{e}(\mathbf{V}^{-1})$ being known, as in (14) [its elements being reciprocals of elements of $\mathbf{e}(\mathbf{V}) = \mathbf{L}_q \boldsymbol{\theta}_q$ given by (11)], and with $\mathbf{L}_q$ having full column rank, (16) yields

$$\boldsymbol{\tau} = \mathbf{L}_q^{+} \mathbf{e}(\mathbf{V}^{-1}) \ , \tag{16}$$

for use in (12).

Q.E.D.

A consequence of this theorem is that under the conditions imposed on the $\mathbf{K}_i$-matrices the inverse of $\mathbf{V} = \Sigma\theta_i \mathbf{K}_i$ can be found more easily than by using the spectral decomposition theorem. Moreover, it derives $\mathbf{V}^{-1}$ in terms of the $\mathbf{K}_i$-matrices that constitute $\mathbf{V}$ itself. It can be derived simply by using $\boldsymbol{\theta}$ and the vectors of eigenroots of the $\mathbf{K}_i$-matrices, which are often more easily obtainable than are those of $\mathbf{V}$ itself, as is necessitated by the spectral decomposition. This is particularly so when the $\mathbf{K}_1$-matrices are patterned matrices, or powers of matrices, or of Kronecker (direct) products of matrices — as is often the case in statistical applications. These are considered in Section 5.

Note that a zero element in $\boldsymbol{\theta}_q$ does not mean that the corresponding τ_i will be zero. Thus a $\mathbf{K}_i$ that does not occur in $\mathbf{V}$ of (1) may occur in $\mathbf{V}^{-1}$ of (12). This is in contrast to Anderson (1969); an example is detailed in Searle and Henderson (1979). Conversely, a non-zero θ_i may have its corresponding τ_i be zero, in which case the corresponding $\mathbf{K}_i$ would occur in $\mathbf{V}$ but not in $\mathbf{V}^{-1}$.

Since $\mathbf{L}_q$ of the set of consistent equations (15) has full column rank, the use of $\mathbf{L}_q^+$ in (17) can be reduced to using a regular inverse, simply by choosing q equations from (15), call them

$$\mathbf{L}_0\boldsymbol{\tau} = \mathbf{e}_0(\mathbf{V}^{-1}) \ , \tag{17}$$

such that $\mathbf{L}_0$ is q linearly independent rows of $\mathbf{L}_q$, and $\mathbf{e}_0(\mathbf{V}^{-1})$ consists of the corresponding elements of $\mathbf{e}(\mathbf{V}^{-1})$. Then

$$\boldsymbol{\tau} = \mathbf{L}_0^{-1}\mathbf{e}_0(\mathbf{V}^{-1}) \ . \tag{18}$$

Finally, we can note that with the complete set of q $\mathbf{K}_i$-matrices being closed under multiplication, $\mathbf{V}$ is then in the commutative quadratic subspace generated by either the idempotent $\mathbf{M}_i$s or by all the q $\mathbf{K}_i$s, and this implies $\mathbf{V}^{-1} = \sum_{i=1}^{q} \tau_i\mathbf{K}_i$ for some τ_i. The spectral decomposition of (3) and (4) is a special case of this. In general, this means that we must always consider all q $\mathbf{K}_i$s that generate the commutative quadratic subspace (i.e., closed under multiplication). This is similar to the quadratic subspace of a vector space of real symmetric matrices introduced by Seely (1971, p. 711). Subsequently, and more appropriately, it has been referred to as a Jordan algebra by Jensen (1975, 1977) and by Seely (1977). The similarity is that Seely's symmetric matrices are replaced by other forms of patterned matrices, e.g., circulants.

Considerable simplification occurs in the special case when one of the **K**-matrices, $\mathbf{K}_1$ say, is **I**. Then

$$\mathbf{V}\mathbf{V}^{-1} = (\Sigma\theta_i\mathbf{K}_i)(\Sigma\tau_i\mathbf{K}_i) = \Sigma\Sigma\theta_i\tau_j\mathbf{K}_i\mathbf{K}_j = \mathbf{I} = \mathbf{K}_1 = \mathbf{K}_1 + \Sigma 0\mathbf{K}_i \ . \qquad (19)$$

Now, closed-under-multiplication gives every $\mathbf{K}_i\mathbf{K}_j$ as a $\mathbf{K}_i$, so the coefficient of every $\mathbf{K}_i$ in $\mathbf{V}\mathbf{V}^{-1}$ is zero, save that of $\mathbf{K}_1$ which is unity. Hence equating coefficients provides linear equations for determining values of the τ_i. This is illustrated at the end of Section 4.1.

4. TWO EXAMPLES WITH SPECIAL PROPERTIES

4.1 A simple dispersion matrix Mukherjee (1976, p. 136) discusses the dispersion matrix (with a > b)

$$\mathbf{V} = \begin{bmatrix} a+b & 0 & a-b \\ 0 & 2a & 0 \\ a-b & 0 & a+b \end{bmatrix} = 2a\begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + 2b\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} , \qquad (20)$$

the latter being the spectral decomposition for eigenroots 2a and 2b with multiplicities 2 and 1, respectively. Hence, from (3)

$$\mathbf{V}^{-1} = \frac{1}{2a}\begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + \frac{1}{2b}\begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} = \frac{1}{4ab}\begin{bmatrix} a+b & 0 & b-a \\ 0 & 2b & 0 \\ b-a & 0 & a+b \end{bmatrix} . \qquad (21)$$

Our theorems are illustrated by writing **V** as

$$\mathbf{V} = (a+b)\mathbf{K}_1 + (a-b)\mathbf{K}_2 \text{ for } \mathbf{K}_1 = \mathbf{I} \text{ and } \mathbf{K}_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} .$$

Then $\mathbf{K}_2^2 = \mathbf{I}$; and $\mathbf{K}_1$ and $\mathbf{K}_2$ are linearly independent, simultaneously diagonalizable, and form a set that is closed under multiplication. Since $\mathbf{e}(\mathbf{K}_1) = [1 \;\; 1 \;\; 1]'$ and $\mathbf{e}(\mathbf{K}_2) = [1 \;\; 1 \;\; -1]'$, Theorem 1 gives

$$\mathbf{e}(\mathbf{V}) = \mathbf{L}_c\boldsymbol{\theta}_c = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a+b \\ a-b \end{bmatrix} = \begin{bmatrix} 2a \\ 2a \\ 2b \end{bmatrix} . \qquad (22)$$

And Theorem 2 gives

$$\boldsymbol{\tau} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1/2a \\ 1/2a \\ 1/2b \end{bmatrix} = \begin{bmatrix} (a+b)/4ab \\ (-a+b)/4ab \end{bmatrix} \tag{23}$$

and so

$$\mathbf{V}^{-1} = \frac{a+b}{4ab}\mathbf{K}_1 + \frac{b-a}{4ab}\mathbf{K}_2 ,$$

which is the same as (21).

The derivation of this arising from $\mathbf{K}_1 = \mathbf{I}$ and given by (19) is to use

$$\mathbf{I} = [(a+b)\mathbf{1} + (a-b)\mathbf{K}_2](\tau_1\mathbf{I} + \tau_2\mathbf{K}_2) .$$

Equating coefficients, using $\mathbf{K}_2^2 = \mathbf{I}$ gives

$$(a+b)\tau_1 + (a-b)\tau_2 = 1 \quad \text{and} \quad (a-b)\tau_1 + (a+b)\tau_2 = 0$$

with the same solution for $\boldsymbol{\tau}$ as in (23).

$\mathbf{L}_c$ in (22) has two rows that are identical, which leads to two elements of $\mathbf{e}(\mathbf{V})$ being equal, i.e., to a multiple root. This characteristic is true more generally: the presence of sets of identical rows in $\mathbf{L}_c$ indicates multiple eigenroots of $\mathbf{V}$. Nevertheless, specifying the multiplicities is difficult because they are not easily related to multiplicities of eigenroots of the $\mathbf{K}_i$-matrices, nor to the ranks of those matrices, their linear independence, or their multiplicative closure. Certainly $\mathbf{L}_c$ has c linearly independent rows, which form a non-singular matrix $\mathbf{L}_*$, and the vector of eigenroots of $\mathbf{V}$ that are linearly independent functions of the elements of $\boldsymbol{\theta}_c$ is then $\mathbf{e}_d(\mathbf{V}) = \mathbf{L}_0\boldsymbol{\theta}_c$, akin to (19). But corresponding multipliities must be obtained as the number of times each row of $\mathbf{L}_0$ occurs in $\mathbf{L}_c$. And even then, these multiplicities pertain only to the eigenroots viewed as linearly independent functions of elements of $\boldsymbol{\theta}_c$. They are not necessarily the multiplicities of the numerical eigenroots, because some of the linearly independent functions of elements of $\boldsymbol{\theta}_c$ may, depending on the numerical value of $\boldsymbol{\theta}_c$, have the same value — and this is what determines multiplicities of numerical eigenroots.

The existence of non-singular $\mathbf{L}_0$ means that the number of eigenroots that are linearly independent functions of θs equals the number of θs. Therefore, just as $\mathbf{e}(\mathbf{V}) = \mathbf{L}_q\boldsymbol{\theta}_q$ of (11) led to $\mathbf{e}(\mathbf{V}^{-1}) = \mathbf{L}_q\boldsymbol{\tau}$ of (16), which gave $\boldsymbol{\tau} = \mathbf{L}_q^+\mathbf{e}(\mathbf{V}^{-1})$ of (16), so now, $\mathbf{e}_d(\mathbf{V}) = \mathbf{L}_0\boldsymbol{\theta}_c$ yields $\boldsymbol{\tau} = \mathbf{L}_0^{-1}\mathbf{e}_d(\mathbf{V}^{-1})$, similar to (18). Thus for the example

$$\mathbf{e}_d(\mathbf{V}) = \mathbf{L}_0\boldsymbol{\theta} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} a+b \\ a-b \end{bmatrix} = \begin{bmatrix} 2a \\ 2b \end{bmatrix},$$

which gives the same result as (23):

$$\boldsymbol{\tau} = \mathbf{L}_0^{-1}\mathbf{e}_d(\mathbf{V}^{-1}) = \tfrac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1/2a \\ 1/2b \end{bmatrix} = \begin{bmatrix} (a+b)/4ab \\ (b-a)/4ab \end{bmatrix}.$$

4.2 Circulants: power structure, and non-symmetry A circulant of order n is a square matrix having the form

$$\mathbf{C} = \mathbf{C}(c_0,\cdots,c_{n-1}) = \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{bmatrix}. \tag{24}$$

It is well known (and easily confirmed) that on defining the "one-element type" circulant $\mathbf{A} = \mathbf{C}(0,1,0,\cdots,0)$, the general circulant $\mathbf{C}$ of (24) can be expressed as

$$\mathbf{C} = \sum_{i=0}^{n-1} c_i\mathbf{A}^i . \tag{25}$$

Since $\mathbf{A}$ and its powers are all permutation matrices (including $\mathbf{A}^0 = \mathbf{I}$), they form a set of linearly independent, simple, commutative (and hence simultaneously diagonizable) matrices that is closed under multiplication. And eigenroots of the one-element type circulant, $\mathbf{A}$, are well-known (e.g., Davis, 1979, p. 73). Application of Theorems 1 and 2 for finding eigenroots of $\mathbf{C}$, and $\mathbf{C}^{-1}$ (if it exists), is therefore quite straight-orward, including being able to use the special case (19) arising from $\mathbf{A}^0 = \mathbf{I}$. This is seen in Gilbert (1962) whose method is just a special case of applying our theorems to circulants. Other methods for obtaining the inverse of a circulant (with the inverse itself being a circulant) are also available in Davis (1979), Searle (1979) and Feinsilver (1984).

Notice that this application of our theorems introduces the idea of having the $\mathbf{K}_i$-matrices of (1) as powers of a matrix — an idea that is extended to Kronecker (direct) products of powers in Section 5. But symmetry is not needed. Nevertheless, symmetric circulants have their place in statistics, as dispersion matrices for certain cyclic partially balanced design, e.g., Wise (1955), Srivastava (1966), Olkin and Press (1969), T.W. Anderson (1969) and D.A. Anderson (1972); but none of these authors appeal to the power structure of (25) that so easily permits using Theorems 1 and 2 for obtaining eigenroots and inverses.

5. EXTENSION TO KRONECKER PRODUCTS

In $\Sigma_i c_i \mathbf{A}^i$ of (25) the general form $\Sigma_i \theta_i \mathbf{K}_i$ of (1) has $\mathbf{K}_i = \mathbf{A}^i$; this is now extended to $\mathbf{K}_i$ being a Kronecker (direct) product of powers of matrices, p of them, say. This is motivated by the matrix

$$\mathbf{V}_p = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{1}} \theta_{\mathbf{i}} \left(\mathbf{J}_p^{i_p} * \mathbf{J}_{p-1}^{i_{p-1}} * \cdots \mathbf{J}_r^{i_r} \cdots * \mathbf{J}_1^{i_1} \right) \qquad (26)$$

(where * represents the Kronecker product operator)[1/] dealt with by Searle

[1/] Footnote for typesetter: * is to be replaced by .

and Henderson (1979) in the context of variance components models. In that case, each $\mathbf{J}_r$ in (26) is square with every element unity, and with each i_r being 0 or 1, which ensures the needed properties of commutativity and multiplicative closure. The linear independent that is also needed arises from each power of a J-matrix in (26) being 0 or 1, with the zero power being an identity matrix. Some (but maybe not all) of the θs are non-zero in variance components models, and represent variances; and some θs may be zero. The exponents in (26), $i_p, \cdots, i_1$, are also used as subscripts to θ where, for notational convenience and readability only, they are represented as a vector $\mathbf{i}$. 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 $r = p, p-1, \cdots, 1$ takes only the values 0 or 1), and the summation $\sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{1}}$ represents the multiple summation $\sum_{i_p=0}^{1} \cdots \sum_{i_1=0}^{1}$ and so notationally is summation over the first 2^p nonnegative binary numbers from $00\cdots0$ to $11\cdots1$.

The generalization of (26) is to have each $\mathbf{K}_i$ of $\Sigma_i \theta_i \mathbf{K}_i$ as a Kronecker product (KP) of a power of each of p matrices $\mathbf{A}_p, \mathbf{A}_{p-1}, \cdots, \mathbf{A}_r, \cdots, \mathbf{A}_1$, with each $\mathbf{A}_r$ being unrelated to the θs. Thus

$$\mathbf{V}_p = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{i}=\mathbf{d}} \theta_{\mathbf{i}} \mathbf{K}_{\mathbf{i}} \tag{27}$$

for

$$\mathbf{K}_{\mathbf{i}} = \mathbf{A}_p^{i_p} * \cdots * \mathbf{A}_r^{i_r} * \cdots * \mathbf{A}_1^{1_i} = \mathop{*}_{r=p}^{r=1} \mathbf{A}_r^{i_r} \tag{28}$$

where in (27) the $\mathbf{i}$ is $[i_p \;\; i_{p-1} \;\cdots\; i_1]'$ and summation is over the range of powers of the $\mathbf{A}_r$-matrices, namely over $\mathbf{i}$ from $\mathbf{0}$ $[0 \;\; 0 \;\cdots\; 0]'$ to $\mathbf{d} = [d_p d_{p-1} \;\cdots\; d_1]'$. When $d_r = d$ for all $r = 1, \cdots, p$, the vector subscript $\mathbf{i}$ will range through the first $(d + 1)^p$ numbers of base $b = d + 1$.

In $\mathbf{K}_i$ of (28), the abbreviated notation on the right-hand side represents the reverse natural order already referred to. Also, to whatever extent not all $\prod_{r=1}^{p}(d+1)$ combinations of powers $0,1,\cdots,d_r$ of A_r occur in V_p, some θs will be zero. For example,

$$\mathbf{V} = \theta_{00}(\mathbf{A}_2^0 * \mathbf{A}_1^0) + \theta_{01}(\mathbf{A}_2^0 * \mathbf{A}_1^1) + \theta_{10}(\mathbf{A}_2^1 * \mathbf{A}_1^0)$$

must be written to additionally include $\theta_{11}(\mathbf{A}_2^1 * \mathbf{A}_1^1)$ so that

$$\mathbf{V}_2 = \mathbf{V} + \theta_{11}(\mathbf{A}_2^1 + \mathbf{A}_1^1) \qquad \text{with } \theta_{11} = 0 \ .$$

And, of course, when $\mathbf{A}_r$ has order $\mathbf{n}_r$, the order of $\mathbf{V}_p$ and of each $\mathbf{K}_i$ is $\mathbf{N}_p = \prod_{r=1}^{p} \mathbf{n}_r$.

We use Theorems 1 and 2 for finding eigenroots and $\mathbf{V}_p^{-1}$, for $\mathbf{V}_p$ of (27). Since eigenroots of $\mathbf{V}_p$ involve those of $\mathbf{K}_i$, as in Theorem 1, they involve the eigenroots of $\mathbf{A}_r$ for $\mathbf{K}_i$ of (28). Thus multiplicities of eigenroots of $\mathbf{V}_p$ occur in abundance, through the multiplicities of eigenroots of powers of the $\mathbf{A}_r$-matrices that occur in Kronecker products in $\mathbf{V}_p$ – see (28). Though (5) – (7), eigenroots of $\mathbf{V}_p$ that are linearly independent functions of the θs can be obtained from manipulating matrices and vectors that often have order considerably less than $\mathbf{V}_p$. All one needs are the distinct eigenroots of $\mathbf{A}_r$, q_r of them, say, and then one deals with matrices of order $\prod_{r=1}^{p} q_r$ rather than of order $\prod_{r=1}^{p}$. This can be a considerable reduction in order on some occasions, particularly so, for example, with $\mathbf{V}_p$ of the variance components models in (26), where a $\mathbf{J}_r$ can have very large order (e.g., $n_r = 500$) but has only two distinct eigenroots, zero with multiplicity $n_r - 1$ and n_r with multiplicity one.

5.1 Eigenroots of V Recall two features of eigenroots. First, those of $\mathbf{A}_r^k$ are the k'th powers of $\mathbf{A}_r$. Second, those of $\mathbf{A} * \mathbf{B}$ are all possible products of eigenroots of **A** and **B**; i.e., $\mathbf{e}(\mathbf{A} * \mathbf{B}) = \mathbf{e}(\mathbf{A}) * \mathbf{e}(\mathbf{B})$. Notice, in passing, how this provides a convenient order, in a vector, of the eigenroots of a Kronecker product:

$$\mathbf{e}(\mathbf{K_i}) = \overset{r=1}{\underset{r=p}{*}} \mathbf{e}(\mathbf{A}_r^{\;i_r}) \; . \tag{29}$$

Then, because each $\mathbf{K_i}$ is a Kronecker-product of powers of the same $\mathbf{A}_r$-matrices, the $\mathbf{K_i}$-matrices are simultaneously diagonalizable, and so Theorem 1 applies directly and gives the vector of eigenroots of $\mathbf{V}_p$ as

$$\mathbf{e}(\mathbf{V}_p) = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{d}} \theta_{\mathbf{i}} \mathbf{e}(\mathbf{K_i}) = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{d}} \theta_{\mathbf{i}} \left[\overset{r=1}{\underset{r=p}{*}} \mathbf{e}(\mathbf{A}_r^{\;i_r}) \right] . \tag{30}$$

Now sequence the values of $\mathbf{i}$ in lexicon order and define $\boldsymbol{\theta}$ similarly,

$$\boldsymbol{\theta} = [\theta_{\mathbf{0}} \cdots \theta_{\mathbf{d}}]' \; .$$

Also, define $\mathbf{L}(\mathbf{A}_r)$ as the $n_r \times (d_r + 1)$ matrix of powers $0,1,2,\cdots,d_r$ of the eigenroots of $\mathbf{A}_r$. Thus for $\mathbf{e}_s(\mathbf{A}_r)$ being the s'th eigenroot of $\mathbf{A}_r$

$$\mathbf{L}(\mathbf{A}_r) = \left\{ [e_s(\mathbf{A}_r)]^{t-1} \right\} \quad \text{for } s = 1,\cdots,n_r \quad \text{and} \quad t = 1,\cdots,d_r+1 \; , \tag{31}$$

where n_r is the order of (square) $\mathbf{A}_r$ and d_r is its highest power in $\mathbf{V}_p$. Then from (30)

$$\mathbf{e}(\mathbf{V}_p) = \mathbf{L}\boldsymbol{\theta} \quad \text{for} \quad \mathbf{L} = \left(\overset{r=1}{\underset{r=p}{*}} [\mathbf{L}(\mathbf{A}_r)] \right) . \tag{32}$$

Example Suppose

$$\mathbf{V} = \theta_{00}(\mathbf{A}^0 * \mathbf{B}^0) + \theta_{01}(\mathbf{A}^0 * \mathbf{B}^1) + \theta_{10}(\mathbf{A}^1 * \mathbf{B}^0)$$

where any one or two of θ_{00}, θ_{01} and θ_{10} may be zero. Then

$$\mathbf{e}(\mathbf{V}) = \theta_{00}\mathbf{e}(\mathbf{A}^0 + \mathbf{B}^0) + \theta_{01}(\mathbf{A}^0 * \mathbf{B}^1) + \theta_{10}(\mathbf{A}^1 + \mathbf{B}^0)$$

$$= \theta_{00}[\mathbf{e}(\mathbf{A}^0) * \mathbf{e}(\mathbf{B}^0)] + \theta_{01}[\mathbf{e}(\mathbf{A}^0) * \mathbf{e}(\mathbf{B}^1)] + \theta_{10}[\mathbf{e}(\mathbf{A}^1) * \mathbf{e}(\mathbf{B}^0)]$$

$$= [\mathbf{e}(\mathbf{A}^0) + \mathbf{e}(\mathbf{B}^0) \quad \mathbf{e}(\mathbf{A}^0) * \mathbf{e}(\mathbf{B}^1) \quad \mathbf{e}(\mathbf{A}^1) * \mathbf{e}(\mathbf{B}^0) \quad \mathbf{e}(\mathbf{A}^1) * \mathbf{e}(\mathbf{B}^1)] \begin{bmatrix} \theta_{00} \\ \theta_{01} \\ \theta_{10} \\ \theta_{11}=0 \end{bmatrix} \tag{33}$$

$$= [\mathbf{e}(\mathbf{A}^0) * \mathbf{e}(\mathbf{A}^1)] * [\mathbf{e}(\mathbf{B}^0) * \mathbf{e}(\mathbf{B}^1)]\boldsymbol{\theta}$$

$$= [\mathbf{L}(\mathbf{A}) * \mathbf{L}(\mathbf{B})]\boldsymbol{\theta} \tag{34}$$

as in (32).

Note how the derivation of (30) from (29) depends upon rewriting **V** so as to include all possible combinations of the powers of (in the example) **A** and **B** that are in fact in **V**, using zero θ-coefficients for the combinations that do not occur in **V**. This is seen in (33) for the example, where $\theta_{11} = 0$ is introduced corresponding to $\mathbf{A}^1 * \mathbf{B}^1$ that does not occur in **V**. Note, too, that the numbers of powers of **A** and **B** need not be the same; e.g., for **A** occurring in **V** with three different powers and **B** with two, there would be six difficult θ-coefficients with maybe some being zero. For **L** to have the Kronecker-product form indicated in (32), it is essential to write $\mathbf{V}_p$ in terms of all $\prod_{r=1}^{p}(d_r + 1)$ $\mathbf{K}_i$-matrices, with zero θ-values where appropriate.

5.2 Inverse of $\mathbf{V}_p$ The preceding result, based as it is on Theorem 1, requires simultaneous diagonability of the $\mathbf{K}_i$-matrices, but not that they be closed under multiplication. But both properties are required for applying Theorem 2 to obtain $\mathbf{V}^{-1}$ when it exists. In order to use (32) in (12) of Theorem 2, the $\mathbf{K}_i$s that need to be included in **V** with zero θs must be not only those needed for having all combinations of the $d_r + 1$ powers of the $\mathbf{A}_r$s with each other [as already specified for **e**(**V**) of (32)], but also those additional $\mathbf{K}_i$s (and possible powers of $\mathbf{A}_r$ greater than d_r) that might be needed for the sake of multiplicative closure. The extent to which this may be necessary will depend upon the actual form of the $\mathbf{A}_r$-matrices. We confine attention to just that closed set of $\mathbf{K}_i$-matrices and let $\mathbf{d}^+ = \left[d_p^+ \cdots d_1^+\right]'$, where $d_r^+ \geq d_r$ is that largest power of $\mathbf{A}_r$ in that closed set.

Example (continued). In (33), we needed $\theta_{11}(\mathbf{A}^1 * \mathbf{B}^1)$ with $\theta_{11} = 0$, in order to have all combinations of powers; but, depending on the exact nature of **A** and **B**, multiplicative closure might also demand including $\theta_{02}(\mathbf{A}^0 * \mathbf{B}^2)$ and $\theta_{12}(\mathbf{A}^1 * \mathbf{B}^2)$ with $\theta_{02} = 0 = \theta_{12}$. Then, whereas $\mathbf{d} = [1 \;\; 1]'$, $\mathbf{d}^+ = [2 \;\; 1]$.

The application of Theorem 2 using (30) – (32) and

$$q = \prod_{r=1}^{r=p} (d_r^+ + 1) \tag{35}$$

is direct. It gives

$$\mathbf{V}_p^{-1} = \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{d}^+} \tau_{\mathbf{i}} \mathbf{K}_{\mathbf{i}} \qquad \text{for } \mathbf{K}_{\mathbf{i}} = \mathop{*}_{r=p}^{r=1} \mathbf{A}_r^{i_r} \tag{36}$$

for

$$\boldsymbol{\tau}_{q\times 1} = (\mathbf{L}_q'\mathbf{L}_q)^{-1}\mathbf{L}_q' \, \mathbf{e}(\mathbf{V}^{-1}) \;, \tag{37}$$

$$\mathbf{L}_q = \mathop{*}_{r=p}^{r=1} \Big[\mathbf{L}(\mathbf{A}_r)\Big]_{n_r\times(d_r^+ +1)} \;, \tag{38}$$

and

$$\mathbf{e}(\mathbf{V}^{-1}) = [1/\delta_1 \;\cdots\; 1/\delta_N], \qquad \text{for } N = \prod_{r=1}^{p} n_r$$

where $\delta_1,\cdots,\delta_N$ are the elements of $\mathbf{e}(\mathbf{V})$ of (32). Thus

$$\boldsymbol{\tau}_q = \mathop{*}_{r=p}^{r=1} (\{[\mathbf{L}(\mathbf{A}_r)]'[\mathbf{L}(\mathbf{A}_r)]\}^{-1}[\mathbf{L}(\mathbf{A}_r)]')\mathbf{e}(\mathbf{V}^{-1}) \;. \tag{39}$$

5.3 A simple example Consider

$$\mathbf{V} = \begin{bmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{bmatrix} \tag{40}$$

$$= a(\mathbf{I} * \mathbf{I}) + b(\mathbf{I} * \mathbf{A}) + c(\mathbf{A} * \mathbf{I}) + d(\mathbf{A} * \mathbf{A})$$

$$= \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{i}=\mathbf{1}} \theta_{\mathbf{i}}(\mathbf{A}^{i_2} * \mathbf{A}^{i_1}) \tag{41}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{with} \quad \mathbf{A}^0 = \mathbf{A}^2 = \mathbf{I} \tag{42}$$

and

$$\boldsymbol{\theta} = [\theta_{00} \quad \theta_{01} \quad \theta_{10} \quad \theta_{11}]' = [a \quad b \quad c \quad d]' \ .$$

In terms of the general notation we here have $p = 2$, $A_2 = A_1 = A$, $n_2 = n_1 = 2$, $d_2 = d_1 = 1 = d_2^+ = d_1^+$ and $q = \Pi(d_r^+ + 1) = 4$.

For obtaining $\mathbf{e}(\mathbf{V})$ from (32) we have from (31) with $e_1(\mathbf{A}) = -1$ and $e_2(\mathbf{A}) = 1$

$$\mathbf{L}(\mathbf{A}_2) = \mathbf{L}(\mathbf{A}_1) = \mathbf{L}(\mathbf{A}) = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \mathbf{L}_2 = \mathbf{L}_1, \text{ say } .$$

Therefore, from (32),

$$\boldsymbol{\delta} = \mathbf{e}(\mathbf{V}) = \mathbf{L}\boldsymbol{\theta} = (\mathbf{L}_2 * \mathbf{L}_1)\boldsymbol{\theta}$$

is

$$\begin{bmatrix} \delta_{00} \\ \delta_{01} \\ \delta_{10} \\ \delta_{11} \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a - b - c + d \\ a + b - c - d \\ a - b + c - d \\ a + b + c + d \end{bmatrix} . \tag{43}$$

Then, with $\mathbf{L}_2 = \mathbf{L}_1$ being non-singular, having inverse

$$\mathbf{L}_2^{-1} = \mathbf{L}_1^{-1} = \tfrac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} ,$$

(37) and (38) gives

$$\boldsymbol{\tau} = (\mathbf{L}_2^{-1} * \mathbf{L}_1^{-1})\mathbf{e}(\mathbf{V}^{-1})$$

as

$$\begin{bmatrix} \tau_{00} \\ \tau_{01} \\ \tau_{10} \\ \tau_{11} \end{bmatrix} = \tfrac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1/\delta_{00} \\ 1/\delta_{01} \\ 1/\delta_{10} \\ 1/\delta_{11} \end{bmatrix} . \tag{44}$$

Then

$$\mathbf{V}^{-1} = \tau_{00}(\mathbf{I} * \mathbf{I}) + \tau_{01}(\mathbf{I} * \mathbf{A}) + \tau_{10}(\mathbf{A} * \mathbf{I}) + \tau_{11}(\mathbf{A} * \mathbf{A})$$

$$= \sum_{\mathbf{i}=\mathbf{0}}^{\mathbf{i}=\mathbf{1}} \tau_{\mathbf{i}}(\mathbf{A}^{i_1} * \mathbf{A}^{i_2}) \ , \tag{45}$$

for the τ-coefficients given by (44).

Of course, since $\mathbf{K}_1 = \mathbf{I} * \mathbf{I} = \mathbf{I}$, the simpler method of equating coefficients based on (19) could also be used. This produces equations $\mathbf{V}\boldsymbol{\tau} = \mathbf{e}_1$ for $\mathbf{e}_1$ being the first column of an identity matrix. Thus $\boldsymbol{\tau}$ is the first column of $\mathbf{V}^{-1}$, which is indeed, also the solution given by (44). Certainly this method is simpler than that used to derive (44); but in this case it needs (the first column of) $\mathbf{V}^{-1}$ in order to get $\mathbf{V}^{-1}$! And, more importantly, that simplicity of equations $\mathbf{V}\boldsymbol{\tau} = \mathbf{e}_1$ does not generalize: those equations arise directly from the precise nature of the multiplicative closure of the set of $\mathbf{K}_i$-matrices, i.e., to which $\mathbf{K}$ is each $\mathbf{K}_i\mathbf{K}_j$ specifically equal.

-i. A correlation matrix. When θ of (41) is $[1 \quad \rho_1 \quad \rho_2 \quad \rho_1]'$, the matrix V in (40) is a quasi-circumplex correlation matrix (Guttman 1954) with eigenvalues coming from (43) as

$$\boldsymbol{\delta} = \mathbf{e}(\mathbf{V}) = [(1-\rho_2) \quad (1-\rho_2) \quad (1-2\rho_1+\rho_2) \quad (1+2\rho_1+\rho_2)]' \ .$$

The resulting spectral decomposition of $\mathbf{V}$ is given by Mukherjee (1976, p. 139 and 1982, p. 448), although without using the Kronecker product structure. A general expression for order $p \times p$ is given in Mukherjee (1984, p. 446).

-ii. A tridiagonal dispersion matrix Putting b = d = 0 in (41) makes it a uniform tridiagonal dispersion matrix that arises in distributed lag models in regression. Mukherjee (1984, p. 444) gives its spectral decomposition as

$$\mathbf{V} = \tfrac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} * \begin{bmatrix} a+c & 0 \\ 0 & a+c \end{bmatrix} + \tfrac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a-c & 0 \\ 0 & a-c \end{bmatrix} .$$

Our formulation has $\mathbf{V}$ as

$$\mathbf{V} = a(\mathbf{I} * \mathbf{I}) + c(\mathbf{A} * \mathbf{I}) \qquad (46)$$

which, with b = d = 0 in (43) gives $\mathbf{e}(\mathbf{V}) = [a-c \quad a-c \quad a+c \quad a+c]'$.

Then (44) gives

$$\boldsymbol{\tau} = (a^2-c^2)^{-1}[a \quad 0 \quad -c \quad 0]'$$

and hence from (45)

$$\mathbf{V}^{-1} = (a^2-c^2)^{-1}[a(\mathbf{I} * \mathbf{I}) - c(\mathbf{A} * \mathbf{I})] ,$$

involving the same Kronecker products as does **V** of (46).

-iii. A need for multiplicative closure When d = 0, there is no $(\mathbf{A} * \mathbf{A})$ term in **V** of (41), but there is such a term in $\mathbf{V}^{-1}$ of (45), thus illustrating the need for having in **V** all terms that constitute multiplicative closure. In fact, in (43), each eigenroot is a linear function of all the θs; and so setting any one θ to zero drops a term from **V** but not from $\mathbf{V}^{-1}$.

5.4 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 $\mathbf{K}_i$ is a Kronecker product of **I**- and **J**-matrices, as in (26). For example, for the two-way crossed classification random model with interaction, the customary dispersion matrix of the vector of observations can be written in the form

$$\mathbf{V} = \sigma_e^2(\mathbf{I}_a * \mathbf{I}_b * \mathbf{I}_n) + \sigma_\alpha^2(\mathbf{I}_a * \mathbf{J}_b * \mathbf{J}_n) + \sigma_\beta^2(\mathbf{J}_a * \mathbf{I}_b * \mathbf{J}_n) + \sigma_{\alpha\beta}^2(\mathbf{I}_a * \mathbf{I}_b * \mathbf{J}_n) .$$

Searle and Henderson (1979) deal with this example (and others) at length, using methods that are essentially just special cases of Theorems 1 and 2 for finding the eigenroots and inverse of **V**, and we do not repeat this example here. Nelder (1979) kindly directed us to his formulating **V** in terms of Kronecker products of **I** and **J** matrices and the consequent spectral decomposition. Nelder (1965a,b) and Thompson (1979) also noted some overlap. Smith and Hocking (1978) also independently develop similar methodology.

Numerous papers have capitalized on these ideas in recent years. Wansbeek and Kapteyn (1982a,b, 1983) and Wansbeek (1985) have considered the spectral decomposition of **V** and Wansbeek (1985) re-analyzes a patterned correlation matrix from Kotz, Pearn and Wichern (1984). Other extensions and analyses are given by Anderson *et al.* (1984), Khuri (1982), Houtman and Speed (1983), Speed (1981, 1983, 1987) and Speed and Bailey (1982), and by Tjur (1984), who has many other references. More recently, Speed (1986, 1987) and Dawid (1987) have taken these ideas to emphasize properties of symmetry in analysis of variance of balanced (equal-subclass-numbers) data. When data are unbalanced the problem of inverting **V** has been considered by Searle and Rudan (1973), Wansbeek (1982) and Bunney and Kissling (1984).

Acknowledgments

Thanks go to referees and to Friedrich Pukelsheim for suggestions that greatly improved the finalizing of this paper, and particularly to the referee who mentioned the special case (19) and its illustration in Section 4.1.

REFERENCES

Anderson, D. A. (1972). Designs with partial factorial balance. *Annals of Mathematical Statistics* **43**, 1333-1341.

Anderson, R. D., Henderson, H. V., Pukelsheim, F. and Searle, S. R. (1984). Best estimation of variance components from balanced data, with arbitrary kurtosis. *Mathematische Operationsforschung und Statistic, Series Statistics* **15**, 163-176.

Anderson, T. W. (1969). Statistical inference for covariance matrices with linear structure. *Multivariate Analysis-II*, Ed., P.R. Krishnaiah, 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.

Anderson, T. W. (1984). *Introduction to Multivariate Statistical Analysis*, 2nd edition, Wiley, New York.

Bonney, G. E. and Kissling, G. E. (1984). On the inverse of a patterned covariance matrix. *Journal of the American Statistical Association* **79**, 722-723.

Davis, Philip J. (1979). *Circulant Matrices*. Wiley, New York.

Dawid, A. P. (1987). Symmetry models and hypotheses for structured data layouts. Research Report #50, Department of Statistical Science, University College London, To appear in *Journal of Royal Statistical Society*.

Feinsilver, Philip (1984). Circulants, inversion of circulants, and some related matrix algebra. *Linear Algebra and Applications* **56**, 29-43.

Gilbert, T. L. (1962). Inverse overlap matrix for periodic arrays of atoms. *Journal of Mathematical Physics* **3**, 107-117.

Guttman, L. (1954). A new approach to factor analysis: the radex. *Mathematical Thinking in the Social Sciences*, ed., P. Lazarsfeld, The Free Press, Glencoe, Illinois, 258-348.

Henderson, Harold V. (1979). Dispersion, patterned and structured matrices: vec operators and related topics. Ph.D. Thesis, Cornell University. 182 pp.

Houtman, A. M. and Speed, T. P. (1983). Balance in designed experiments with orthogonal block structure. *Annals of Statistics* **11**, 1069-1085.

Jensen, S. T. (1975). Covariance hypotheses which are linear in both the covariance and the inverse covariance. Preprint No. 1, Institute of Mathematical Statistics, University of Copenhagen.

Jensen, S. T. (1977). An algebraic theory for normal statistical models with unknown covariance. *Recent Developments in Statistics*, ed, J.R. Barra *et al.*, North-Holland.

Khuri, A. I. (1982). Direct products: a powerful tool for the analysis of balanced data. *Communications in Statistics, A: Theory and Methods* **11**, 2903-2920.

Kotz, S., Pearn, W. L. and Wichern, D. W. (1984). Eigenvalue-eigenvector analysis for a class of patterned correlation matrices with an application. *Statistics and Probability Letters* **2**, 119-125.

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, J. J. (1977). Asymptotic properties of maximum likelihood estimates in the mixed model of the analysis of variance. *Annals of Statistics* **5**, 746-762.

Mirsky, L. (1955, 1982). *An Introduction to Linear Algebra*, Clarendon, Oxford, 1955, and Dover, New York, 1982.

Mukherjee, B. N. (1976). Techniques of covariance structural analysis. *Australian Journal of Statistics* **18**, 131-150.

Mukherjee, B. N. (1982). Measures of multivariate relationships for the uniform covariance case. *Behaviormetrika* **12**, 27-66.

Mukherjee, B. N. (1984). A simple approach to testing hypotheses regarding a class of covariance structures. *Statistics: Applications and New Directions*, eds., J.K. Ghosh and J. Roy, Indian Statistical Institute, Calcutta.

Nelder, J. A. (1965a). The analysis of randomized experiments with orthogonal block structure, I. Block structure and the null analysis of variance. *Proceedings of the Royal Statistical Society (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. *Proceedings of the Royal Statistical Society (London), Ser. A* **273**, 163-178.

Nelder, J. A. (1979). Personal Communication.

Olkin, I. and Press, S. J. (1969). Testing and estimation for a circular stationary model. *Annals of Mathematical Statistics* **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 Applications* **25**, 77-89.

Searle, S. R. and Henderson, H. V. (1979). Dispersion matrices for variance component models. *Journal of the American Statistical Association* **74**, 465-470.

Searle, S. R. and Rudan, J. W. (1973). Wanted: an inverse matrix. *Communications in Statistics* **2**, 155-166.

Seber, G. A. F. (1984). *Multivariate Observations*, Wiley, New York.

Seely, Justus (1971). Quadratic subspaces and completeness. *Annals of Mathematical Statistics* **42**, 710-721.

Seely, Justus (1977). Minimal sufficient statistics and completeness for multivariate normal families. *Sankhyā, Series A* **39**, 170-185.

Smith, David W. and Hocking, R. R. (1978). Maximum likelihood analysis of the mixed mocel: the balanced case. *Communications in Statistics, A: Theory and Methods* **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*, eds. S. Kotz and N.L. Johnson, Vol. **3**, Wiley, New York.

Speed, T. P. (1986). ANOVA models with random effects: an approach via symmetry. *Essays in Time Series and Allied Processes: Papers in Honor of E.J. Hannan*, eds. J. Gani and M.B. Priestly. *Journal of Applied Probability, Special Volume* **23A**, 355-368.

Speed, T. P. (1987). What is an analysis of variance? *Annals of Statistics* **15**, 885-941.

Speed, T. P. and Bailey, R. A. (1982). On a class of association schemes derived from lattices of equivalence relations. *Algebraic Structures and Applications*, eds. sP. Schultz, C.E. Praeger and R.P. Sullivan. Marcel Dekker, New York, 55-74.

Srivastava, J. N. (1966). On testing of hypotheses regarding a class of covariance structures. *Psychometrica* **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. *Annals of Statistics* **8**, 802-810.

Szatrowski, Ted H. (1980). Necessary and sufficient conditions for explicit solutions in the multivariate normal estimation problem for patterned means and covariances. *Annals of Statistics* **8**, 802-810.

Szatrowski, T. H. (1985). Patterned covariances. *Encyclopedia of Statistical Science*, eds. S. Kotz and N.L. Johnson, Vol. **6**, Wiley, New York, 638-641.

Szatrowski, Ted H. and Miller, John J. (1980). Explicit maximum likelihood estimates from balanced data in the mixed model of the analysis of variance. *Annals of Statistics* **8**, 811-819.

Takeuchi, K., Yanai, H. and Mukherjee, B. N. (1983). *The Foundations of Multivariate Analysis: A Unified Approach by Means of Projection onto Linear Subspaces*, Halsted Press, New York.

Thompson, R. (1979). Review 419.62061 of Searle and Henderson (1979). *Zentralblatt fur Mathematik*.

Tjur, T. (1984). Analysis of variance models in orthogonal designs. *International Statistical Review* **52**, 33-81.

Wansbeek, Tom (1982). Another approach to inverting a covariance matrix when data are unbalanced. *Communications in Statistics, Simulation and Computation* **11**, 583-588.

Wansbeek, Tom (1985). Eigenvalue-eigenvector analysis for a class of patterned correlation matrices with an application: a comment. *Statistics and Probability Letters* **3**, 95-96.

Wansbeek, Tom and Kapteyn, Arie (1982a). A class of decompositions of the variance-covariance matrix of a generalized error components model. *Econometrics* **50**, 713-724.

Wansbeek, Tom and Kapteyn, Arie (1982b). A simple way to obtain the spectral decompositon of variance components models for balanced data. *Communications in Statistics, A: Theory and Methods* **11**, 2105-2112.

Wansbeek, Tom and Kapteyn, Arie (1983). A note on spectral decomposition and maximum likelihood estimation in ANOVA models with balanced data. *Statistics and Probability Letters* **1**, 213-215.

Wise, J. (1955). The autocorrelation function and the spectral density function. *Biometrika* **42**, 151-159.