privileged communication

ON A CLASS OF PATTERNED MATRICES
-INTERIM REPORT-

MARLOS VIANA

Abstract. The covariance matrix of the ordered equicorrelated normal distribution characterizes the normal distribution and is a member of a class of covariance matrices with group structure. The study of its subgroups leads to patterned matrices of special interest such as matrix circulants.

1. Introduction

Patterned covariance matrices arise from a variety of contexts. Wilks (1946), in one of the early papers with patterned structure, considered a set of measurements on k equivalent psychological tests. This led to a covariance matrix with equal diagonal elements and equal off-diagonal elements. Votaw (1948) extended this model to a set of blocks in which each block had a pattern. Olkin (1973) studied a multivariate version in which each element was a matrix, and the blocks were patterned.

The circular covariance matrix has a long history. There, typical measurements arise from a physical, spatial or temporal model, as for example, measurements on the petals of a flower. For a discussion of this model see Olkin and Press (1969) and Olkin (1973).

The structure of patterned matrices arising in statistics was studied by a number of authors. Roy and Sarhan (1956) obtained the inverse of matrices that are function of a Green's matrix $G = (g_{ij}),\ g_{ij} = a_i b_j,\ i < j$, e.g., Horn and Johnson (1991, p. 508). Other matrices considered are those studied by Votaw (1948).

Roy, Greenberg and Sarhan (1960) considered matrices of the form

$$D_a + \alpha \mathbf{v}\mathbf{v}',$$

where $D_a = \operatorname{diag}(a_1, \ldots, a_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$. The covariance matrix of the multinomial distribution is of this form. Graybill (1969, Chapter 8) summarizes a number of results on patterned matrices.

Greenberg and Sarhan (1959) are motivated by linear model analyses in which an inverse is required. In some cases, such as response-surface analysis, observations are taken in a manner that generates a patterned covariance matrix.

Statistics has motivated a number of papers. Barrett and Feinsilver (1978) give a probabilistic proof that the inverse of a tridiagonal matrix (a Jacobi matrix) is a Green's matrix. Their results is based on a Gaussian Markovian family.

Date: August 29, 2003.

1991 *Mathematics Subject Classification.* nnn-mmm.

Key words and phrases. Covariance Matrices, Circulants, Order Statistics.

943b.

In a series of papers (we give the latest which provides references to earlier papers), Szatrowski (1985) discusses how to obtain maximum likelihood estimates for the elements of a class of patterned covariance matrices. Chinchilli and Carter (1984) considered a patterned covariance arising from a multivariate growth-curve model. Browne (1977) reviews patterned correlation matrices arising from multiple psychological measurements. This paper provides references to related papers. In particular, the work of Guttman (1954) introduces a number of patterns.

Anderson (1969) and Rogers and Young (1977) studied the case in which a covariance matrix Σ can be written as $\Sigma = \theta_1 G_1 + \ldots + \theta_n G_n$, where the θ's are unknown parameters and $G_1, \ldots G_n$ is a set of symmetric linearly independent matrices. The intraclass covariance matrix and the circular covariance matrix fall into this class. When the matrices $G_1, \ldots, G_n$ are simultaneously diagonalizable (by an orthogonal matrix independent of the θ's), then explicit maximum likelihood estimates are readily obtainable.

Of course we cannot expect many patterns to have linear structure, and some patterns arising from certain group operations will not have linear structure. Diaconis (1989) provides a discussion of how such patterns can arise. These matrices that are exhibited are

$$A = \begin{bmatrix} a & b & b & b & b & c \\ b & a & b & c & b & b \\ b & b & a & b & c & b \\ b & c & b & a & b & b \\ b & b & c & b & a & b \\ c & b & b & b & b & a \end{bmatrix},$$

$$B = \begin{bmatrix} a & b & c & d & e \\ c & d & e & a & b \\ e & a & b & c & e \\ b & c & d & e & a \\ d & e & a & b & c \end{bmatrix}, \quad C = \begin{bmatrix} a & b & b & c & c & b \\ b & a & c & b & b & c \\ b & c & a & b & b & c \\ c & b & b & a & c & b \\ c & b & b & c & a & b \\ b & c & c & b & b & a \end{bmatrix}.$$

Of these, A and B have linear structure, whereas C does not. A simple determinant of linear structure is to show that G_i and G_j commute, or equivalently, that $G_i G_j$ is symmetric.

The class of patterned matrices considered in the following sections arise in several ways. Let the components of $\mathbf{U} = (U_0, U_1, \ldots, U_n)$ be independent and identically distributed real random variables with mean zero and unit variance and α_0 and α_1 be fixed real numbers. The random vector $\mathbf{Y}$ with components

$$Y_j = \alpha_0 U_0 + \alpha_1 U_j, \quad j = 1, \ldots, n, \tag{1.1}$$

has an equicorrelated covariance matrix of the form

$$\text{Cov}(Y) = a_1 \mathbf{e}\mathbf{e}' + a_2 \mathbf{I}, \tag{1.2}$$

where $\mathbf{e}$ is the $n \times 1$ vector of ones, $a_1 = \alpha_0^2$ and $a_2 = \alpha_1^2$. The common non-negative correlation among the components of $\mathbf{Y}$ is

$$\rho = \frac{a_1}{a_1 + a_2}.$$

The covariance matrix of the corresponding vector $\mathcal{Y}$ of ordered observations $(Y_{(1)}, \ldots, Y_{(n)})$, $Y_{(1)} \leq \ldots \leq Y_{(n)}$, is (mod orientation)

$$\mathrm{Cov}(\mathcal{Y}) = a_1 \mathbf{e}\mathbf{e}' + a_2 A, \tag{1.3}$$

where A is the covariance matrix of the ordered components of $\mathbf{U}$ (Olkin and Viana 1995).

When the components of $\mathbf{U}$ are independent and normally distributed with mean 0 and variance 1 then $A = (a_{ij})$ is a symmetric stochastic matrix, that is, $\sum_j a_{ij} = \sum_i a_{ij} = 1$, for $i, j = 1, \ldots, n$ (David 1981, p. 39). The values of a_{ij} are tabulated, e.g., Sarhan and Greenberg (1962); Owen (1962).

Note that the equicorrelated matrix (1.2) has the symmetry of S_n, the group of all permutation matrices of dimension n, in the sense that it commutes with every member S_n. Conversely, if a matrix M commutes with every member of S_n then M has the pattern of (1.2). In fact, if a matrix M commutes with every member of S_n then M commutes with their sum, which is (proportional to) $\mathbf{e}\mathbf{e}'$. The roots of $\mathbf{e}\mathbf{e}'$ are 1 with multiplicity one and eigenvector proportional to $\mathbf{e}$, and 0 with multiplicity $n-1$ and eigenvectors $\mathbf{v}$ orthogonal to $\mathbf{e}$. Because M and $\mathbf{e}\mathbf{e}'$ are real symmetric commuting matrices, the resulting spectral decomposition M has the form $\alpha_1 \mathbf{e}\mathbf{e}' + \alpha_2 \mathbf{v}\mathbf{v}'$, for some real numbers α_1 and α_2. In addition, because $\mathbf{I} = \mathbf{e}\mathbf{e}' + (n-1)\mathbf{v}\mathbf{v}'$, it follows that $M = a_1 \mathbf{e}\mathbf{e}' + a_2 \mathbf{I}$, for some real numbers a_1 and a_2.

On the other hand, the (doubly) stochasticity of A in (1.3) characterizes the normal distribution. Let $Z_1, \ldots, Z_n$ be independent, identically distributed random variables with zero means and variance σ^2, $Z_{(1)} \leq \ldots \leq Z_{(n)}$ its ordered values and $\mathcal{Z} = (Z_{(1)}, \ldots, Z_{(n)})$. Then $\mathrm{Cov}(\mathcal{Z}) = \mathcal{C}$ is doubly stochastic if and only if $Z_1, \ldots, Z_n$ are normal, for all $n = 2, 3, \ldots$.

The proof of doubly stochasticity for normal random variables is provided in David (1981, p. 39). Suppose, conversely, that $\mathcal{C}$ is doubly stochastic. Then

$$1 = \sum_{j=1}^{n} \mathrm{Cov}(Z_{(i)}, Z_{(j)}) = \mathrm{Cov}(Z_{(i)}, \sum_{j=1}^{n} Z_{(j)}) = \mathrm{Cov}(Z_{(i)}, n\bar{Z}),$$

$i = 1, \ldots, n$ and $n = 2, \ldots$. But this is a characterization of the normal distribution, (Govindarajulu 1966) [see also Johnson and Kotz (1970, p. 53)].

Algebraic operations among matrices with the structure of (1.3) may arise as follows: Start with a non-null vector $\mathbf{v}_0$ of dimension n and define

$$\mathbf{v}_k = (a_{1,k} \mathbf{e}\mathbf{e}' + a_{2,k} A)\mathbf{v}_{k-1} \equiv M_k \mathbf{v}_{k-1}, \quad k = 1, \ldots,$$

where the sequence $(a_{1,k}, a_{2,k})$ may be selected according to a random distribution and is such that $na_{1,k} + a_{2,k} = 1$ at each step of k. Then,

$$\mathbf{v}_k = M_k M_{k-1} \ldots M_1 \mathbf{v}_0,$$

and the sequence $\{\mathbf{v}_0, \mathbf{v}_1, \ldots\}$ is a random walk on the plane orthogonal to the equiangular line and distant the average value of the components of $\mathbf{v}_0$ from the origin. At each step, the components of $\mathbf{v}_k$ are convex combinations of their simple average and weighted averages of $\mathbf{v}$ with weights defined by the rows of A.

2. THE CLASS $a_1\mathbf{e}\mathbf{e}' + a_2A$.

Let $GL(n,\mathbb{R})$ indicate the multiplicative group of $n \times n$ real non-singular matrices and

$$G = \{A \in GL(n,\mathbb{R}); A\mathbf{e} = \mathbf{e}, \mathbf{e}'A = \mathbf{e}'\}$$

be the set of all non-singular $n \times n$ real doubly-stochastic matrices. Given $A \in G$ and a real vector $a = (a_1, a_2)$ define the $n \times n$ matrix

$$[a, A] = a_1\mathbf{e}\mathbf{e}' + a_2A, \tag{2.1}$$

where $\mathbf{e}$ is the n-component vector of ones. It then follows that

$$[a, A][b, B] = (na_1b_1 + a_1b_2 + a_2b_1)\mathbf{e}\mathbf{e}' + a_2b_2AB, \tag{2.2}$$

suggesting the operation

$$* : (a, b) \in \mathbb{R}^2 \times \mathbb{R}^2 \to a * b = (na_1b_1 + a_1b_2 + a_2b_1,\ a_2b_2) \in \mathbb{R}^2, \tag{2.3}$$

so that

$$[a, A][b, B] = [a * b, AB]. \tag{2.4}$$

Proposition 2.1. $(\mathbb{R}^2, +, *)$ is an algebra.

Proof. In fact, for all $a, b, c \in \mathbb{R}^2$ and all numbers γ in the scalar field ($\mathbb{R}$ or $\mathbb{C}$) of the vector space $(\mathbb{R}^2, +)$, we have

$$\begin{aligned} a * b &\in (\mathbb{R}^2, +), \\ a * (b + c) &= a * b + a * c, \\ (a + b) * c &= a * c + b * c, \\ \gamma(a * b) &= a * (\gamma b) = (\gamma a) * b. \end{aligned}$$

□

Proposition 2.2. $W = \{(a_1, a_2) \in \mathbb{R}^2;\ \ a_2 \neq 0,\ \ na_1 + a_2 \neq 0\}$ together with $*$ is a commutative group.

Proof. Let $a, b \in W$ and $a*b = (c_1, c_2)$. Then $c_2 = a_2b_2 \neq 0$ and $nc_1+c_2 = (na_1+a_2)(nb_1+b_2) \neq 0$, so that $a * b \in W$. The unit of $(W, *)$ is $1_* = (0, 1) \in W$. Now let, for $a \in W$,

$$a_*^{-1} = (\frac{-a_1}{a_2(na_1 + a_2)}, \frac{1}{a_2}).$$

By virtue of the definition of W, a_*^{-1} is well defined, $a * a_*^{-1} = a_*^{-1} * a = 1_*$ and $a_*^{-1} \in W$ (to see this, note that if $a_*^{-1} = (c_1, c_2)$ then both $nc_1 + c_2 = 1/(na_1 + a_2)$ and $c_2 = 1/a_2$ are non-zero). The associativity and commutativity of (2.3) can be verified directly. □

Example 2.1. $(\{(-2/n, 1), (0, 1)\}, *)$ is a subgroup of $(W, *)$ of order 2.

Example 2.2. Let $\mathbb{Z}_n$ indicate the integers under addition and multiplication modulo (prime) n. Then $na_1 + a_2 = a_2$, $W = \{(a_1, a_2) \in \mathbb{Z}_n^2; a_2 \neq 0\}$, which together with the multiplication $*$ of (2.3) taken modulo n, is an Abelian group of order $(n-1)n$ isomorphic to $C_{n-1} \times C_n$. Take $n = 3$: The 6 members of W are

$$\begin{aligned} 1 &= (0, 1), \quad a = (0, 2), \\ b &= (1, 1), \quad c = (1, 2), \\ d &= (2, 1), \quad e = (2, 2). \end{aligned}$$

The resulting multiplication table

$*$	1	a	b	c	d	e
1	1	a	b	c	d	e
a	a	1	e	d	c	b
b	b	e	d	a	1	c
c	c	d	a	b	e	1
d	d	c	1	e	b	a
e	e	b	c	1	a	d

shows that W is isomorphic to

$$\{\mathbf{I}, (12)(36)(45), (135)(264), (143256), (153)(246), (165234)\} \subset S_6,$$

via $g \to \pi_g$, with $\pi_g(h) = g * h,\ g, h \in W$. The element a is of order 2, b and d have order 3, e and c have order 6. W is isomorphic to $C_2 \times C_3$ via

$$a^\alpha b^\beta \to (a^\alpha, b^\beta).$$

The irreducible representations have the form

$$a^\alpha b^\beta \to w_a^\alpha w_b^\beta, \quad w_a^2 = 1, w_b^3 = 1.$$

There are 6 of these, determined by $w_a = 1, -1$ and $w_b = 1, w, \bar{w}$. The associated character table is

	1	a	b	c	d	e
χ_1	1	1	1	1	1	1
χ_2	1	1	w	w^2	w^2	w
χ_3	1	1	v	v^2	v^2	v
χ_4	1	-1	1	-1	1	-1
χ_5	1	-1	w	$-w_2$	w^2	$-w$
χ_6	1	-1	v	$-v_2$	v^2	$-v$

, $v = \bar{w}$.

Proposition 2.3. $WG = \{[a, A]; a \in W, A \in G\}$ is a subgroup of $GL(n, \mathbb{R})$.

Proof. First note that G is a subgroup of $GL(n, \mathbb{R})$. The direct product group $W \times G$ of W and G is defined by the operation

$$(a, A), (b, B) \quad \to \quad (a * b, AB),$$

which turns $(a, A), (b, B) \to [a * b, AB]$ into an homomorphism between $W \times G$ and $GL(n, \mathbb{R})$ such that $(1_*, 1_G) = ((0, 1), \mathbf{I}_n) \ \to\ \mathbf{I}_n$. Therefore WG is an isomorphic image of the product group and hence a subgroup of $GL(n, \mathbb{R})$. ☐

Example 2.3. $G = \{\mathbf{I}_n\}$ generates the subgroup of all equicorrelated covariance matrices; $W = \{1_*\}$ and $G = S_n$ generate the subgroup S_n of $n \times n$ permutation matrices; $W = \{1_*\}$ and

$$G = \{w_0\mathbf{I}_n + w_1 g + w_2 g^2 + \ldots + w_{n-1} g^{n-1};\ \sum_{i=0}^{n-1} w_i = 1,\ w_i \in \mathbb{R}\},$$

where g is an element of order n in S_n, generate the subgroup of stochastic circulants with first row $\mathbf{w}' = (w_0, \ldots, w_{n-1})$. Take $n = 4$ and let F be a stochastic circulant with first row $\mathbf{w}'$. Then $F' = w_0\mathbf{I} + w_1 g^3 + w_2 g^2 + w_3 g \in G$ and

$$FF' = \alpha_0\mathbf{I} + \alpha_1 g + \alpha_2 g^2 + \alpha_1 g^3 = \begin{bmatrix} \alpha_0 & \alpha_1 & \alpha_2 & \alpha_1 \\ \alpha_1 & \alpha_0 & \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_1 & \alpha_0 & \alpha_1 \\ \alpha_1 & \alpha_2 & \alpha_1 & \alpha_0 \end{bmatrix}$$

is a symmetric stochastic circulant with first row determined by $\alpha_i = \mathbf{w}' g^i \mathbf{w}$.

Example 2.4. The covariance matrix $A \in G$ of n ordered independent standard normal components indicated in (1.3) has a doubly symmetric stochastic structure with the additional property that $A_{i,j} = A_{n+1-j,n+1-i}$. In this example we look at properties of such matrices. When $n = 4$,

$$A = \begin{bmatrix} h_1 & a & b & c \\ a & h_2 & d & b \\ b & d & h_2 & a \\ c & b & a & h_1 \end{bmatrix} = \begin{bmatrix} .491 & .245 & .158 & .104 \\ .245 & .360 & .235 & .158 \\ .158 & .235 & .360 & .245 \\ .104 & .158 & .245 & .491 \end{bmatrix}$$

can be expressed as

$$A = \beta \mathbf{I} + a[(12)(34)] + b[(13)(24)] + c[(14)] + d[(23)], \tag{2.5}$$

where $[(ij)]$ and $[(ij)(kl)]$ indicate the corresponding permutation representation of permutations (ij) and $(ij)(kl)$. Expression (2.5) implies that

$$h_1 = \beta + d, \quad h_2 = \beta + c,$$

whereas the stochastic property $A\mathbf{e} = \mathbf{e}$ requires that

$$\beta + a + b + c + d = 1.$$

The quantity β is an invariant when A is transformed by gAg', $g \in S_4$. When $n = 4$ the invariant corresponds to

$$\begin{aligned} \beta &= \mathrm{Var}(Y_{(1)}) - \mathrm{Cov}(Y_{(2)}, Y_{(3)}), \\ &= \mathrm{Var}(Y_{(2)}) - \mathrm{Cov}(Y_{(1)}, Y_{(4)}). \end{aligned}$$

The following are examples of expression (2.5) and corresponding invariants indexed for $n = 2, 3, 5, 6$:

$$\begin{aligned} A_2 &= \beta_2 \mathbf{I} + a[(12)], \\ \beta_2 &= h_1, \\ A_3 &= \beta_3 \mathbf{I} + a\{[(12)] + [(23)]\} + b[(13)], \\ \beta_3 &= h_1 - a = h_2 - b, \\ A_5 &= \beta_5 \mathbf{I} + a[(12)(45)] + b\{[(13)] + [(35)]\} + c[(14)(25)] \\ &\quad + d[(15)] + e\{[(23)] + [(34)]\} + f[(24)], \\ \beta_5 &= h_1 - b - 2e - f = h_2 - 2b - d - e = h_3 - a - c - d - f, \\ A_6 &= \beta_6 \mathbf{I} + a[(12)(56)] + b[(13)(46)] + c[(14)(36)] + d[(15)(26)] + \\ &\quad + e[(16)] + f[(23)(45)] + g[(24)(35)] + h[(25)] + i[(34)], \\ \beta_6 &= h_1 - f - g - h - i = h_2 - b - c - e - i = h_3 - a - d - e - h. \end{aligned}$$

We note that A is a linear combination of single transpositions (2-cycles) or products of two transpositions, and that all transpositions are present. Because A is a stochastic matrix, Birkhoff's Theorem anticipates that A is necessarily a linear combination of permutation matrices. The presence of all transpositions reflects the fact that 2-cycles alone are sufficient to sort n objects. In general, we have

$$A_n = \beta_n \mathbf{I} + \sum_{i<j\leq n} a_{ij} M(i,j),$$

where, for $i < j \leq n$,

$$M(i,j) = \begin{cases} [(i\ j)(n+1-j\ \ n+1-i)] & \text{when } j \neq n+1-j, \\ [(i\ j)] + [(n+1-j\ \ n+1-i)] & \text{when } j = n+1-j. \end{cases} \tag{2.6}$$

The resulting $m = n/2$ conjugate-invariant relations among the entries of A_n when n is even and $m = (n+1)/2$ relations when n is odd are obtained from

$$h_t = \beta_n + \sum_{i<j\leq n} a_{ij}\delta_t(i,j), \quad 1 \leq t \leq m, \tag{2.7}$$

where

$$\delta_t(i,j) = \begin{cases} 1 - \delta_{\{i,j,n+1-j,n+1-i\}}(t), & \text{when } j \neq n+1-j, \\ 2 - \delta_{\{i,j\}}(t) - \delta_{\{n+1-j,n+1-i\}}(t), & \text{when } j = n+1-j, \end{cases}$$

and δ_L is the indicator function of set L. The cases illustrated above also show that the permutation matrices associated to distinct coefficients are symmetric. However, they do not jointly commute and do not form a subgroup of the underlying symmetric group. It can also be shown that the class of A matrices considered in this example is not closed under multiplication. However, if A is non-singular, and has a doubly symmetric stochastic structure satisfying $A_{i,j} = A_{n+1-j,n+1-i}$ then so does its inverse.

The largest subgroup of S_n fixing A has order 2, e.g., $\{\mathbf{I}, g_n\}$:

$$g_2 = (12),\ g_3 = (13),\ g_4 = (14)(23),\ g_5 = (15)(24),\ g_6 = (16)(25)(34).$$

Because g_n has order 2 and trace 1 when n is odd and trace 0 when n is even it follows that g_n has an eigenvalue 1 with multiplicity $m_1 = n/2$ and an eigenvalue -1 with multiplicity $m_2 = n/2$, when n is even. Otherwise $m_1 = (n+1)/2$ and $m_2 = n - m_1$. Therefore, A factors into $A = A_+ \oplus A_-$, where the dimension of A_+ is m_1^2 and A_- has dimension m_2^2. Moreover, A_+ has a one-dimensional factor associated with an eigenvector proportional to $\mathbf{e}$, so that the factorization

$$A = \{1\} \oplus B_+ \oplus A_-,$$

obtains, where $\{1\}$ of dimension one is the constant unit sum of the row (or column) entries of A, B_+ is symmetric with dimension $(m_1 - 1)^2$ and A_- is symmetric with dimension m_2^2. This, in turn, leads to the factorization of $[a, A] = a_1\mathbf{e}\mathbf{e}' + a_2 A$ into

$$[a, A] = \{na_1 + a_2\} \oplus a_2 B_+ \oplus a_2 A_-.$$

Proposition 2.4. For every member $[a, A]$ of WG, the following properties hold:

(1) $\text{tr}([a, A]) = na_1 + \text{tr}(A)a_2$,
(2) $[a, A]/(na_1 + a_2)$ is doubly stochastic,
(3) $\mathbf{e}'[a, A]\mathbf{e} = n(na_1 + a_2)$, the sum of all entries of $[a, A]$,
(4) $[a, A] = [a, \mathbf{I}_n][1, A]$,
(5) Denote by λ the diagonal matrix of eigenvalues. Then $\lambda([a, \mathbf{I}_n]) = \text{diag}(na_1 + a_2, a_2, \ldots, a_2)$,
(6) $\det[a, A] = (na_1 + a_2)a_2^{n-1}\det A$,
(7) If A is symmetric, $\lambda([a, A]) = \lambda([a, \mathbf{I}_n])\lambda([1, A])$,
(8) $[a, A]^{-1} = [a^{-1}, A^{-1}]$,
(9) $[a, A]^k = [a^k, A^k]$,
(10) Define $[a, A]_\otimes = a_1 \otimes \mathbf{e}\mathbf{e}' + a_2 \otimes A$. Then $[a, A] \otimes [b, A] = [a * b, A^2]_\otimes$.

Proposition 2.5. For all $[a, A]$ in WG we have

$$e^{[a,A]} = \frac{1}{n}(e^{na_1} - 1)e^{a_2}\mathbf{e}\mathbf{e}' + e^{a_2 A}.$$

If A is symmetric with eigenvalues $1, \lambda_2, \ldots, \lambda_n$, then

$$\lambda(e^{[a,A]}) = \operatorname{diag}(e^{na_1+a_2}, e^{a_2\lambda_2} \ldots, e^{a_2\lambda_n}).$$

Proof. Because the matrices $\mathbf{e}\mathbf{e}'$ and A commute, we have

$$e^{[a,A]} = e^{a_1\mathbf{e}\mathbf{e}' + a_2 A} = e^{a_1\mathbf{e}\mathbf{e}'}e^{a_2 A}.$$

Also,

$$e^{a_1\mathbf{e}\mathbf{e}'} = \sum_{k\geq 0} \frac{(a_1\mathbf{e}\mathbf{e}')^k}{k!} = \mathbf{I} + \frac{1}{n}(e^{na_1} - 1)\mathbf{e}\mathbf{e}'.$$

The first part of the proposition then follows from the fact that $\mathbf{e}\mathbf{e}'e^{a_2 A} = \mathbf{e}\mathbf{e}'e^{a_2}$. To obtain the corresponding eigenvalues, suppose that A is symmetric, so that

$$Qe^{[a,A]}Q' = e^{Q[a,A]Q'}.$$

The result follows $\lambda([a, A]) = \lambda([a, \mathbf{I}_n])\lambda([1_*, A])$, as shown above on Proposition 2.4. □

From proposition 2.5 we obtain,

$$e^{[a,A]}\mathbf{e} = e^{na_1+a_2}\mathbf{e}, \;\; \mathbf{e}'e^{[a,A]} = e^{na_1+a_2}\mathbf{e}'$$

showing that $e^{[a,A]}$ when divided by $e^{na_1+a_2}$ is a doubly stochastic matrix, and hence a member of WG. The function

$$\gamma(u) = \frac{e^{[a,A]u}}{e^{(na_1+a_2)u}} \in WG, \quad 0 <\mid u \mid\leq d, \; d > 0 \tag{2.8}$$

plays an important part in determining the dimension of WG, defined as the dimension of its tangent space WG_T [e.g., Curtis (1984)]. Vectors in the subspace WG_T are tangent vectors $\Gamma'(0)$ to curves $\Gamma : (-d, d) \rightarrow WG$ such that $\Gamma(0) = \mathbf{I}$. In fact, note that $\gamma(0) = \mathbf{I}$, whereas direct computation shows that

$$\gamma'(0) = [a, A] - (na_1 + a_2)\mathbf{I}.$$

In particular, taking $a_1 = 0$, $a_2 = 1$ and A to be a permutation matrix $g \in G$ we obtain that WG_T includes all matrices $\gamma_g^{'}(0) = g - \mathbf{I}$. Because these matrices span a subspace of dimension $n(n-1)$ we have the following result:

Proposition 2.6. $n(n - 1) \leq \dim WG \leq n^2$.

3. THE GENERAL CLASS $a_1\mathbf{e}\mathbf{e}' + a_2 A$.

A closer look at Proposition 2.2 shows that what makes it work is the fact that $(\mathbb{R}, +, .)$ is a ring and $\mathbb{R} - \{0\}$ is a multiplicative group. Let R be a ring with unit 1_R and L a multiplicative subgroup of $GL(n, R)$ (a division sub ring would also work). Redefine, for $a, b \in GL(n, R)^2$,

$$* : (a, b) \rightarrow a * b = ((\mathbf{e}'\mathbf{e})a_1b_1 + a_1b_2 + a_2b_1, \; a_2b_2) \in GL(n, R)^2, \tag{3.1}$$

where $\mathbf{e}'\mathbf{e} = 1_R + \ldots + 1_R$. Note that $\mathbf{e}\mathbf{e}'$ is an element in the center of R. Let also

$$W_R = \{(a_1, a_2) \in GL(n, R)^2; \; (\mathbf{e}'\mathbf{e})a_1 + a_2 \in L, \; a_2 \in L\},$$

$$G_R = \{A \in GL(n, R); A\mathbf{e} = \mathbf{e}, \mathbf{e}'A = \mathbf{e}'\},$$

$$WG_R = \{[a, A]; a \in W_R, A \in G_R\}.$$

Proposition 3.1. W_R together with $*$ of (3.1) is a group, whereas WG_R is a multiplicative subgroup of $GL(n,R)$. If, in addition, R is commutative, then so is W_R.

Proof. Direct verification. ☐

Similar result hold when, given $\mathbf{x} \in R^n$,

$$\begin{aligned} W_{R,x} &= \{(a_1,a_2) \in GL(n,R)^2;\ (\mathbf{x}'\mathbf{x})a_1 + a_2 \in L,\ a_2 \in L - \{0\}\}, \\ G_{R,x} &= \{A \in GL(n,R); A\mathbf{x} = \mathbf{x},\ \ \mathbf{x}'A = \mathbf{x}'\}, \\ WG_{R,x} &= \{a_1\mathbf{x}\mathbf{x}' + a_2A;\ a \in W_{R,x}, A \in G_{R,x}\}. \end{aligned}$$

Proposition 3.2. If $\mathbf{x}'\mathbf{x}$ is an element in the center of R then $W_{R,x}$ together with

$$* : (a,b) \to a * b = ((\mathbf{x}'\mathbf{x})a_1b_1 + a_1b_2 + a_2b_1,\ a_2b_2) \in GL(n,R)^2$$

is a group, whereas $WG_{R,x}$ is a multiplicative subgroup of $GL(n,R)$. If, in addition, R is commutative, then so is $W_{R,x}$.

Proof. Direct verification. ☐

Example 3.1. The class $\otimes WG = \{a_1 \otimes \mathbf{e}\mathbf{e}' + a_2 \otimes A; a \in W_R, A \in G_R\}$. In $GL(2,\mathbb{R})$, consider the multiplicative subgroup,

$$L = \{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\}.$$

The members of $W = \{(a_1,a_2) \in GL(2,\mathbb{R})^2;\ 2a_1 + a_2 \in L,\ a_2 \in L\}$ are

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

$$c = (\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}),\ \ d = (0, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}),$$

whereas the non-singular doubly stochastic matrices have the form

$$A(t) = \begin{bmatrix} t & 1-t \\ 1-t & t \end{bmatrix},\ \ t \neq \frac{1}{2}.$$

Let $G = L = \{A(1), A(0)\}$, instead of the whole group of doubly stochastic matrices. The resulting members of $\otimes WG$ are:

$$1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ \ a = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

$$b = \frac{1}{2}\begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix},\ \ c = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix},$$

$$d = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix},\ \ e = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix},$$

$$f = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad g = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

with commutative multiplication table

$*$	1	a	b	c	d	e	f	g
1	1	a	b	c	d	e	f	g
a	a	1	c	b	e	d	g	f
b	b	c	1	a	f	g	d	e
c	c	b	a	1	g	f	e	d
d	d	e	f	g	1	a	b	c
e	e	d	g	f	a	1	c	b
f	f	g	d	e	b	c	1	a
g	g	f	e	d	c	b	a	1

.

$\otimes WG$ is isomorphic to the subgroup with members in S_8,

$$\mathbf{I}, (12)(34)(56)(78),\ (13)(24)(57)(68),\ (14)(23)(58)(67),$$
$$(15)(26)(37)(48),\ (16)(25)(38)(47),\ (17)(28)(35)(46),\ (18)(27)(36)(45)$$

all of which have order 2, and to $C_2 \times C_2 \times C_2$ via

$$a^\alpha c^\gamma e^\epsilon \to (a^\alpha, c^\gamma, e^\epsilon).$$

The irreducible representations have the form

$$a^\alpha c^\gamma e^\epsilon \to w_a^\alpha w_c^\gamma w_e^\epsilon, \quad w_a^2 = 1, w_c^2 = 1, w_e^2 = 1.$$

There are 8 of these, determined by $w_a, w_c, w_e \in \{1, -1\}$. The associated character table is

	1	a	b	c	d	e	f	g
χ_1	1	1	1	1	1	1	1	1
χ_2	1	1	1	1	−1	−1	−1	−1
χ_3	1	1	−1	−1	1	1	−1	−1
χ_4	1	−1	−1	1	−1	1	1	−1
χ_5	1	1	−1	−1	−1	−1	1	1
χ_6	1	−1	−1	1	1	−1	−1	1
χ_7	1	−1	1	−1	−1	1	−1	1
χ_8	1	−1	1	−1	1	−1	1	−1

.

Proposition 3.3. Let $[a, A] \in \otimes WG$ over a ring R. Then

1. $[a, A] = [a, \mathbf{I}][1_*, A]$;
2. $(\Gamma \otimes \mathbf{I}_n)[a, \mathbf{I}](\Gamma' \otimes \mathbf{I}_n) = \text{block-diag}(na_1 + a_2, a_2, \ldots, a_2)$, where Γ is any orthogonal matrix such that $\Gamma \mathbf{e}\mathbf{e}' \Gamma' = \text{diag}(n, 0, \ldots, 0)$.
3. $\det[a, I] = \det(na_1 + a_2)\det^{n-1}(a_2)$,
4. If a_1 and a_2 are symmetric commuting matrices, then

$$\det[a, \mathbf{I}_n] = \prod_{j=1}^{n} (n\alpha_{1j} + \alpha_{2j})\alpha_{2j}^{n-1},$$

where $\alpha_{i1}, \ldots, \alpha_{in}$ are the eigenvalues of a_i, $i = 1, 2$.

Proof. Parts 1 and 4 are direct verification. Parts 2 and 3 follow from observing that part 5 of Proposition 2.4 holds for W over an arbitrary ring R with unit 1. $\square$

The University of Illinois at Chicago
Colleges of Medicine and LAS,
1855 W. Taylor Street EEI (m/c 648)
Chicago IL 60612
viana@uic.edu

REFERENCES

Anderson, T. W. (1969), Statistical inference for covariance matrices with linear structure, *in* P. R. Krishnaiah, ed., 'Multivatiate Analysis II', Academic Press, New York, pp. 55–66.

Barrett, W. W. and Feinsilver, P. J. (1978), 'Gaussian families and a theorem of patterned matrices', *Journal of Applied Probability* **15**, 514–522.

Browne, M. W. (1977), 'The analysis of patterned correlation matrices by generalized least squares', *British Journal of Mathematical and Statistical Psychology* **30**, 113–124.

Chinchilli, V. M. and Carter, W. (1984), 'A likelihood ratio test for a patterned covarianve matrix in a multivariate growth-curve model', *Biometrics* **40**, 151–156.

Curtis, M. L., ed. (1984), *Matrix Groups*, Springer-Verlag, New York, NY.

David, H. A. (1981), *Order Statistics*, Wiley, New York.

Diaconis, P. (1989), Patterned matrices, Technical Report 320, Stanford University Department of Statistics, Stanford, California.

Govindarajulu, Z. (1966), 'Characterization of normal and generalized truncated normal distributions using order statistics', *The Annals of Mathematical Statistics* **37**, 1011–1015.

Graybill, F. A. (1969), *Matrices with Applications in Statistics*, Wadsworth, Belmont, California.

Greenberg, B. G. and Sarhan, A. E. (1959), 'Matrix inversion, its interest and application in analysis of data', *Journal of the American Statistical Association* **54**, 755–766.

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

Horn, R. A. and Johnson, C. R. (1991), *Topics in Matrix Analysis*, Cambridge University Press, Cambridge.

Johnson, N. L. and Kotz, S. (1970), *Distributions in Statistics: Continuous Univariate Distributions - 1*, John Wiley and Sons, New York.

Olkin, I. (1973), Testing and estimation for structures which are circularly symmetric in blocks, *in* D. G. Kabe and R. P. Gupta, eds, 'Multivariate Statistical Inference', North-Holland, Amsterdam, pp. 183–195.

Olkin, I. and Press, S. (1969), 'Testing and estimation for a circular stationary model', *The Annals of Mathematical Statistics* **40**, 1358–1373.

Olkin, I. and Viana, M. A. G. (1995), 'Correlation analysis of extreme observations from a multivariate normal distribution', *Journal of the American Statistical Association* **90**, 1373–1379.

Owen, D. B. (1962), *Handbook of Statistical Tables*, Addison-Wesley, Reading, MA.

Rogers, G. and Young, D. (1977), 'Explicit maximum likelihood estimators for certain patterned covariance matrices', *Communications in Statistics - Theory and Methods* **A6**(2), 121–133.

Roy, S. N. and Sarhan, E. (1956), 'On inverting a class of patterned matrices', *Biometrika* **43**, 227–231.

Roy, S. N., Greenberg, B. G. and Sarhan, A. E. (1960), 'Evaluation of determinants, characteristic equations and their roots for a class of patterned matrices', *Journal name to confirm* **00**, 348–359.

Sarhan, A. E. and Greenberg, B. G. (1962), *Contributions to Order Statistics*, Wiley, New York.

Szatrowski, T. H. (1985), 'Asymptopic distributions in the testing and estimation of the missing-data multivariate normal linear patterned mean and correlation matrix', *Linear Algebra and its Applications* **67**, 215–231.

Votaw, D. F. (1948), 'Testing compound symmetry in a normal multivariate distribution', *The Annals of Mathematical Statistics* **19**, 447–473.

Wilks, S. S. (1946), 'Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution', *Annals of Mathematical Statistics* **17**, 257–281.