A Characterization of Matrix Groups That Act Transitively on the Cone of Positive Definite Matrices*

Steen A. Andersson
Department of Mathematics
Indiana University
Bloomington, Indiana 47405

and

Michael D. Perlman
Department of Statistics
University of Washington
Seattle, Washington 98195

Submitted by Friedrich Pukelsheim

Dedicated to Ingram Olkin

ABSTRACT

It is well known that the group of all nonsingular lower block-triangular $p \times p$ matrices acts transitively on the cone $\mathscr{P}^*$ of all positive definite $p \times p$ matrices. This result has been applied to obtain several major results in multivariate statistical distribution theory and decision theory. Here a converse is established: if a matrix group acts transitively on $\mathscr{P}^*$, then its group algebra must be (similar to) the algebra of all lower block-triangular $p \times p$ matrices with respect to a fixed partitioning. This implies the nonexistence of multivariate normal linear statistical models with unrestricted covariance structure that admit a transitive group action, other than those classical models invariant under a Full block-triangular group.

1. INTRODUCTION

It is well known that the group $\mathscr{T}$ of all nonsingular lower triangular $p \times p$ matrices acts *transitively* on the cone $\mathscr{P}^*$ of all positive definite

*Research supported in part by the Danish Research Council, by U.S. National Science Foundation Grant 89-02211, and by U.S. National Security Agency Grant MDA 904-92-H-3083.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

$p \times p$ matrices, or equivalently,[1]

$$\mathscr{P}^* = \mathscr{T}\mathscr{T}^t := \{ AA^t | A \in \mathscr{T} \}.$$

This fact often has been exploited to obtain major results in multivariate statistical distribution theory and decision theory. For example, a classical derivation of the distribution of a random Wishart matrix $S \sim W_p(\Sigma, n)$ is based on the representation $S = AA^t$, where $A \in \mathscr{T}$ (cf. Anderson, 1984, Chapter 7). James and Stein (1961) and Olkin and Selliah (1977) used this transitive action to construct estimators of the covariance matrix Σ which uniformly dominate the classical estimator $n^{-1}S$; together with the Hunt-Stein theorem, this demonstrates the inadmissibility and nonminimaxity of $n^{-1}S$. Giri, Kiefer, and Stein (1963) used this transitive action to establish the minimaxity of Hotelling's T^2 test.

Of course, any group consisting of all nonsingular *block*-triangular $p \times p$ matrices (with respect to a fixed partitioning) also acts transitively on $\mathscr{P}^*$; this extended fact has been used to study the decision-theoretic and distributional properties of many other multivariate normal models and testing problems that remain invariant under such groups. These include the MANOVA and generalized MANOVA problems (Anderson, 1984, Chapter 8; Marden, 1983), testing problems for means with covariates (Giri, 1968; Marden and Perlman, 1980), missing- or additional-data models (Eaton and Kariya, 1983, and stepdown procedures (Marden and Perlman, 1990)—see Andersson, Marden, and Perlman (1994) for a unified treatment of such problems. More examples and references appear in Giri (1977) and Eaton (1983).

Because of the statistical importance of these transitive actions, a natural question arises: *are there any matrix groups other than the full block-triangular groups that act transitively on* $\mathscr{P}^*$? The answer to this question as stated is trivially yes. For example, the proper subgroup $\mathscr{T}^+ \subset \mathscr{T}$ consisting of all lower triangular $p \times p$ matrices with positive diagonal elements also acts transitively on $\mathscr{P}^*$. However, *the groups* $\mathscr{T}$ *and* $\mathscr{T}^+$ *span the same matrix algebra*, i.e., $\mathrm{Alg}(\mathscr{T}) = \mathrm{Alg}(\mathscr{T}^+) =$ the algebra of *all* lower triangular $p \times p$ matrices. Furthermore, in any multivariate normal linear model[2] the invariance group $\mathscr{G}$ is presented in the form $\mathscr{G} = \mathscr{A}(\mathscr{U})^*$, the set of all nonsingular matrices in an algebra $\mathscr{A}(\mathscr{U})$ [see (2.6)] determined by a set of linear constraints. Therefore we are led to the following reformulated question: *if* $\mathscr{G}$ *is a matrix group that acts transitively on* $\mathscr{P}^*$, *must* $\mathrm{Alg}(\mathscr{G})$ *be a (generalized) block-triangular matrix algebra*? (See Definition 2.7.)

[1] Here, t denotes "transpose."

[2] Such as those referenced in the preceding paragraph.

If the answer to this revised question were no, then such groups $\mathscr{G}$ would determine new multivariate normal models (i.e., those that remain invariant under such groups) with unrestricted covariance structure yet with tractable decision-theoretic and distributional properties. Perhaps unfortunately, however, our main result (Theorem 3.1) answers this revised question affirmatively: *Any matrix algebra $\mathscr{A}$ containing a matrix group that acts transitively on $\mathscr{P}^*$ must be a generalized block-triangular algebra of $p \times p$ matrices.* Thus: *For a multivariate normal linear statistical model with unrestricted covariance structure the assumption of transitivity does not allow the appearance of invariance groups essentially different than the classical block-triangular groups.*

This result is used by Andersson, Marden, and Perlman (1994) to characterize totally ordered multivariate normal linear models, i.e., those models that impose no restriction on the covariance structure and that remain invariant under some full block-triangular matrix group. Such models appear to be the only multivariate normal linear models with unrestricted covariance structure that admit explicit (noniterative) maximum-likelihood estimators and likelihood-ratio tests. It follows from our main result that a multivariate normal linear model is totally ordered if and only if the group of all model-preserving linear transformations acts transitively on the model.

After some preliminary results regarding transitive action and block-triangular matrices in Section 2, the main results are presented in Section 3, followed by the proofs of two key lemmas in Sections 4 and 5. All vector spaces and matrices considered in this paper are real, but the main results remain valid (with the obvious modifications) in the complex case[3] where the cone $\mathscr{P}^*$ of all real positive definite symmetric matrices is replaced by the cone of all complex positive definite Hermitian matrices.

2. PRELIMINARIES: TRANSITIVE ACTION AND BLOCK-TRIANGULAR MATRICES

It will be notationally convenient to work with vectors and matrices having unordered index sets. For any two finite index sets I and J, let $\mathscr{M}(I \times J)$ denote the set of all $I \times J$ matrices with real entries, and let $\mathscr{M}(I) := \mathscr{M}(I \times$

[3] In fact, some of the proofs are easier in the complex case—for example, the proof of Lemma 3.2 given in Section 4.

I). For any subsets $\mathscr{A} \subseteq \mathscr{A}(I)$ and $U \subseteq \mathbb{R}^I$ and fixed $E, F \in \mathscr{A}(I)$, define

$$\mathscr{A}^* := \{A \in \mathscr{A} \mid A \text{ is nonsingular}\},$$

$$\mathscr{A}U := \{Ax \mid A \in \mathscr{A},\ x \in U\},$$

$$E\mathscr{A}F := \{EAF \mid A \in \mathscr{A}\}.$$

For any subset $\mathscr{A} \subseteq \mathscr{A}(I)^*$ define

$$\mathscr{A}^{-1} = \{A^{-1} \mid A \in \mathscr{A}\}.$$

A subset $\mathscr{A} \subseteq \mathscr{A}(I)$ is a *matrix algebra* if $\mathscr{A}$ is closed under addition, scalar multiplication, and matrix multiplication. We shall only consider algebras $\mathscr{A}$ that also contain the $I \times I$ identity matrix 1_I. Clearly $\mathscr{A}(I)$ and $\{\lambda 1_I \mid \lambda \in \mathbb{R}\}$ are the maximal and minimal such algebras in $\mathscr{A}(I)$. If $\mathscr{A}$ is an algebra, then[4] $(\mathscr{A}^*)^{-1} = \mathscr{A}^*$, so $\mathscr{A}^*$ is a group under matrix multiplication. For any subset $\mathscr{A} \subseteq \mathscr{A}(I)$ let $\mathrm{Alg}(\mathscr{A})$ denote the algebra generated by $\mathscr{A}$, i.e., the smallest algebra in $\mathscr{A}(I)$ that contains $\mathscr{A}$ and 1_I. If $\mathscr{A} \subseteq \mathscr{A}(I)$ is an algebra, then[5] $\mathrm{Alg}(\mathscr{A}^*) = \mathscr{A}$.

Let $\mathscr{P}(I)$ [or $\mathscr{P}(I)^*$] denote the cone of all positive semidefinite [or positive definite] $I \times I$ matrices. For any subset $\mathscr{A} \subseteq \mathscr{A}(I)$ [or $\mathscr{A}(I)^*$] and $\Sigma \in \mathscr{P}(I)$ [or $\mathscr{P}(I)^*$] define

$$\mathscr{A}\mathscr{A}^t := \{AA^t \mid A \in \mathscr{A}\} \subseteq \mathscr{P}(I) \ [\text{or } \mathscr{P}(I)^*],$$

$$\mathscr{A}\Sigma\mathscr{A}^t := \{A\Sigma A^t \mid A \in \mathscr{A}\} \subseteq \mathscr{P}(I) \ [\text{or } \mathscr{P}(I)^*],$$

and similarly define $\mathscr{A}^t\mathscr{A}$ and $\mathscr{A}^t\Sigma\mathscr{A}$.

DEFINITION 2.1. A group $\mathscr{G} \subseteq \mathscr{A}(I)^*$ acts *transitively* on $\mathscr{P}(I)^*$ if $\mathscr{G}\Sigma\mathscr{G}^t = \mathscr{P}(I)^*$ for every $\Sigma \in \mathscr{P}(I)^*$.

---

[4] It suffices to show that $A \in \mathscr{A}^* \Rightarrow A^{-1} \in \mathscr{A}^*$. Let $f(\lambda) \equiv \det(A - \lambda 1_I)$ be the characteristic polynomial of A, having degree $|I|$. Then $f(0) = \det A \neq 0$, while $f(A) = 0$ by the Cayley-Hamilton theorem. Thus $A^{-1} = \{A^{-1}[f(0)1_I - f(A)]\}/f(0)$, but this is a polynomial of degree $|I| - 1$ in A; hence $A^{-1} \in \mathscr{A}^*$.

[5] The inclusion $\mathrm{Alg}(\mathscr{A}^*) \subseteq \mathscr{A}$ is trivial. If $A \in \mathscr{A}$, let $f(\lambda)$ be the characteristic polynomial of A, and choose λ such that $f(\lambda) \neq 0$. Then $A - \lambda 1_I$ is nonsingular; hence $A \equiv (A - \lambda 1_I) + \lambda 1_I \in \mathrm{Alg}(\mathscr{A}^*)$.

REMARK 2.2. If $\mathscr{G}_1$ acts transitively on $\mathscr{P}(I)^*$ and $\mathscr{G}_1 \subseteq \mathscr{G}_2$, then $\mathscr{G}_2$ also acts transitively on $\mathscr{P}(I)^*$.

When $\mathscr{G} = \mathscr{A}^*$ for some matrix algebra $\mathscr{A}$, the following proposition relates the transitivity of $\mathscr{A}^*$ to similar conditions on $\mathscr{A}$.

PROPOSITION 2.3. *For any matrix algebra $\mathscr{A} \subseteq \mathscr{A}(I)$, the following ten conditions are equivalent*:

(i) $\mathscr{A}\mathscr{A}^t = \mathscr{P}(I)$.
(ii) $\mathscr{A}^*(\mathscr{A}^*)^t = \mathscr{P}(I)^*$.
(iii) $\mathscr{A}\Sigma\mathscr{A}^t = \mathscr{P}(I)^*$ *for some* $\Sigma \in \mathscr{P}(I)^*$.
(iv) $\mathscr{A}^*\Sigma(\mathscr{A}^*)^t = \mathscr{P}(I)^*$ *for some* $\Sigma \in \mathscr{P}(I)^*$.
(v) $\mathscr{A}\Sigma\mathscr{A}^t = \mathscr{P}(I)$ *for every* $\Sigma \in \mathscr{P}(I)^*$.
(vi) $\mathscr{A}^*\Sigma(\mathscr{A}^*)^t = \mathscr{P}(I)^*$ *for every* $\Sigma \in \mathscr{P}(I)^*$ *(i.e.,* $\mathscr{A}^*$ *acts transitively on* $\mathscr{P}(I)^*$*).*
(vii) $E\mathscr{A}F(E\mathscr{A}F)^t = \mathscr{P}(I)$ *for some pair* $E, F \in \mathscr{A}(I)^*$.
(viii) $E\mathscr{A}^*F(E\mathscr{A}^*F)^t = \mathscr{P}(I)^*$ *for some pair* $E, F \in \mathscr{A}(I)^*$.
(ix) $E\mathscr{A}F(E\mathscr{A}F)^t = \mathscr{P}(I)$ *for every pair* $E, F \in \mathscr{A}(I)^*$.
(x) $E\mathscr{A}^*F(E\mathscr{A}^*F)^t = \mathscr{P}(I)^*$ *for every pair* $E, F \in \mathscr{A}(I)^*$.

Furthermore, these ten conditions are equivalent to each of the ten additional conditions (i′)–(x′) obtained by interchanging $\mathscr{A}$ and $\mathscr{A}^t$, $\mathscr{A}^*$ and $(\mathscr{A}^*)^t$, $E\mathscr{A}F$ and $(E\mathscr{A}F)^t$, and $E\mathscr{A}^*F$ and $(E\mathscr{A}^*F)^t$.

Proof. The implications (i) ⇒ (ii), (iii) ⇒ (iv), (v) ⇒ (vi), (vii) ⇒ (viii), and (ix) ⇒ (x) are immediate. To show that (ii) ⇒ (i), note that for any $\Omega \in \mathscr{P}(I)$ there is a sequence $\{\Omega_n\} \subseteq \mathscr{P}(I)^*$ such that $\Omega_n \to \Omega$. By (ii), $\exists\{A_n\} \subseteq \mathscr{A}^*$ such that $A_n A_n^t = \Omega_n$. Since $\{A_n\}$ is bounded, there is a convergent subsequence $\{A_{n'}\}$ such that $A_{n'} \to A \in \mathscr{A}$ (since $\mathscr{A}$, being a finite-dimensional vector space, is closed). By continuity, $AA^t = \Omega$. The proofs that (iv) ⇒ (iii), (vi) ⇒ (v), (viii) ⇒ (vii), and (x) ⇒ (ix) are similar.

The implication (ix) ⇒ (vii) is trivial, while (vii) ⇒ (iii) is immediate, since $\mathscr{A}(I)^*[\mathscr{A}(I)^*]^t = \mathscr{P}(I)^*$ and $\mathscr{P}(I) = E\mathscr{P}(I)E^t$. To show that (iii) ⇒ (i), choose $A \in \mathscr{A}$ such that $A\Sigma A^t = 1_I$. (Necessarily, $A \in \mathscr{A}^*$ and $A^{-1} \in \mathscr{A}^*$). Thus $\mathscr{P}(I) = \mathscr{A}A^{-1}(A^{-1})^t\mathscr{A}^t = \mathscr{A}A^{-1}(\mathscr{A}A^{-1})^t = \mathscr{A}\mathscr{A}^t$, since $\mathscr{A}$ is an algebra. To show that (i) ⇒ (v), for any $\Sigma \in \mathscr{P}(I)^*$ choose $A \in \mathscr{A}$ such that $AA^t = \Sigma$. (Again, $A \in \mathscr{A}^*$ and $A^{-1} \in A^*$). Then $\mathscr{P}(I) = \mathscr{A}A^{-1}\Sigma(A^{-1})^t\mathscr{A}^t = \mathscr{A}A^{-1}\Sigma(\mathscr{A}A^{-1})^t = \mathscr{A}\Sigma\mathscr{A}^t$ as before. Finally, to show that (v) ⇒ (ix), for every pair $E, F \in \mathscr{M}(I)^*$ we have that $\mathscr{P}(I) = \mathscr{A}FF^t\mathscr{A}^t$ and $\mathscr{P}(I) = E\mathscr{P}(I)E^t$; hence $\mathscr{P}(I) = E\mathscr{A}FF^t\mathscr{A}^tE^t = E\mathscr{A}F(E\mathscr{A}F)^t$.

The equivalence of (i′)–(x′) is proved analogously. Lastly, to show that (ii) ⇔ (ii′) $(\mathscr{A}^*)^t\mathscr{A}^* = \mathscr{P}(I)^*$, just note that $[\mathscr{A}^*(\mathscr{A}^*)^t]^{-1} = [(\mathscr{A}^*)^{-1}]^t(\mathscr{A}^*)^{-1} = (\mathscr{A}^*)^t\mathscr{A}^*$ and $\mathscr{P}(I)^* = [\mathscr{P}(I)^*]^{-1}$. ■

We now introduce the algebras of block-triangular matrices. For any subset $K \subseteq I$ and $x \equiv (x_i \mid i \in I) \in \mathbb{R}^I$ let $x_K := (x_i \mid i \in K)$ denote the coordinate projection[6] of x onto $\mathbb{R}^K$. Define the linear subspace $U_K \subseteq \mathbb{R}^I$ as follows:

$$U_K = \{x \in \mathbb{R}^I \mid x_K = 0\}; \tag{2.1}$$

note that[7] $K \subset K' \Rightarrow U_k \supset U_{K'}$. For any $A \in \mathscr{A}(I)$ let A_K denote the $K \times K$ submatrix of A.

Let $\mathscr{D}(I)$ denote the set of all subsets of I. For any set $\mathscr{K} \subseteq \mathscr{D}(I)$ define

$$\mathscr{A}(\mathscr{K}) := \{A \in \mathscr{A}(I) \mid \forall K \in \mathscr{K},\ AU_K \subseteq U_K\}. \tag{2.2}$$

Clearly $\mathscr{A}(\mathscr{K})$ is a matrix subalgebra of $\mathscr{A}(I)$, and for any $\mathscr{K}_1, \mathscr{K}_2 \subseteq \mathscr{D}(I)$,

$$\mathscr{K}_1 \subseteq \mathscr{K}_2 \quad \Rightarrow \quad \mathscr{A}(\mathscr{K}_2) \subseteq \mathscr{A}(\mathscr{K}_1). \tag{2.3}$$

A set $\mathscr{K} \subseteq \mathscr{D}(I)$ is called a *ring* if it is closed under $\cap$ and $\cup$ and if $\varnothing, I \in \mathscr{K}$.[8] For $K\ (\neq \varnothing) \in \mathscr{K}$, define

$$\langle K \rangle := \bigcup (L \in \mathscr{K} \mid L \subset K) \subseteq K,$$

$$[K] := K \setminus \langle K \rangle,$$

$$J(\mathscr{K}) := \{K \in \mathscr{K} \mid K \neq \varnothing, [K] \neq \varnothing\};$$

$J(\mathscr{K})$ is the set of *join-irreducible* elements of $\mathscr{K}$. Then[9] for each $K \in J(\mathscr{K})$,

$$K = \dot{\bigcup}([L] \mid L \in J(\mathscr{K}), L \subseteq K), \tag{2.4}$$

where $\dot{\bigcup}$ denotes a disjoint union.

For any $A \in \mathscr{A}(I)$ and any two subsets $K, L \in J(\mathscr{K})$, let $A_{[KL]}$ denote the $[K] \times [L]$ submatrix of A.

---

[6] Define $x_\varnothing = 0$; thus $U_\varnothing = \mathbb{R}^I$.

[7] In this paper, $\subset$ and $\supset$ are used to indicate *strict* inclusion.

[8] For any $\mathscr{K} \subseteq \mathscr{D}(I)$, $\mathscr{A}(\mathscr{K}) = \mathscr{A}(\text{Ring}(\mathscr{K}))$, where $\text{Ring}(\mathscr{K})$ is the ring generated by $\mathscr{K}$. Thus when studying $\mathscr{A}(\mathscr{K})$, we may always assume that $\mathscr{K}$ is a ring.

[9] This is well known; e.g., see Andersson and Perlman (1993, Proposition 2.1).

PROPOSITION 2.4. *Let $\mathscr{K}$ be a ring of subsets of I. The following three conditions on $A \in \mathscr{A}[I]$ are equivalent:*

(i) $A \in \mathscr{A}(\mathscr{K})$.
(ii) $\forall x \in \mathbb{R}^I$, $\forall K \in \mathscr{K}$, $(Ax)_K = A_K x_K$.
(iii) $\forall K, L \in J(\mathscr{K})$, $L \not\subseteq K \Rightarrow A_{[KL]} = 0$.

Proof. Since

$$\mathscr{A}(\mathscr{K}) = \{A \in \mathscr{A}(I) \,|\, \forall K \in \mathscr{K},\ x_K = 0 \Rightarrow (Ax)_K = 0\},$$

the implication (ii) ⇒ (i) is trivial, while (iii) ⇒ (ii) follows from (2.4) and the usual formula for block matrix multiplication:

$$(Ax)_K = \left(\sum \left(A_{[LM]} x_{[M]} \,\middle|\, M \in J(\mathscr{K})\right) \,\middle|\, L \in J(\mathscr{K}),\ L \subseteq K\right)$$

$$= \left(\sum \left(A_{[LM]} x_{[M]} \,\middle|\, M \in J(\mathscr{K}),\ M \subseteq K\right) \,\middle|\, L \in J(\mathscr{K}),\ L \subseteq K\right)$$

[by (iii)]

$$= A_K x_K.$$

To show that (i) ⇒ (iii), consider $K, L \in J(\mathscr{K})$ with $L \not\subseteq K$. Then for any $x \in \mathbb{R}^I$ such that $x_{[M]} = 0\ \forall M \in J(\mathscr{K})$, $M \neq L$,

$$(Ax)_{[K]} = \sum \left(A_{[KM]} x_{[M]} \,\middle|\, M \in J(\mathscr{K})\right) = A_{[KL]} x_{[L]}.$$

But $(Ax)_K = 0$ by (i); hence $(Ax)_{[K]} = 0$. Since $x_{[L]}$ is arbitrary, $A_{[KL]} = 0$. ■

A ring $\mathscr{K}$ is a *chain* if it is totally ordered[10] under inclusion (hence finite); in this case $J(\mathscr{K}) = \mathscr{K} \setminus \{\varnothing\}$. The equivalence of (i) and (iii) in Proposition 2.4 leads to the following definition:

DEFINITION 2.5. Let $\mathscr{K}$ be a chain of subsets of I. The algebra $\mathscr{A}(\mathscr{K})$ is called the *algebra of block-triangular matrices with respect to $\mathscr{K}$*. The group $\mathscr{A}(\mathscr{K})^*$ is called the *group of block-triangular matrices with respect to $\mathscr{K}$*.

[10] That is, for any distinct $K, L \in \mathscr{K}$ either $K \subset L$ or $L \subset K$.

EXAMPLE 2.6. If $\mathscr{K} = \{\emptyset, I\}$ then $\mathscr{A}(\mathscr{K}) = \mathscr{A}(I)$. If $\mathscr{K} = \{\emptyset, K, I\}$ where $\emptyset \subset K \subset I$, then by Proposition 2.4(iii),

$$\mathscr{A}(\mathscr{K}) = \left\{ \begin{pmatrix} A_{[KK]} & 0 \\ A_{[IK]} & A_{[II]} \end{pmatrix} \in \mathscr{A}(I) \,\middle|\, A_{[KK]} \in \mathscr{A}(K), \right.$$

$$\left. A_{[IK]} \in \mathscr{A}((I \setminus K) \times K),\ A_{[II]} \in \mathscr{A}(I \setminus K) \right\},$$

where the matrices are partitioned according to the decomposition $I = [K] \dot{\cup} [I] \equiv K \dot{\cup} (I \setminus K)$ [cf. (2.4)].

If $\mathscr{K}_1$ and $\mathscr{K}_2$ are chains, then $\mathscr{K}_1 \subset \mathscr{K}_2 \Rightarrow J(\mathscr{K}_1) \subset J(\mathscr{K}_2)$, so by Proposition 2.4(iii), (2.3) can be sharpened as follows:

$$\mathscr{K}_1 \subset \mathscr{K}_2 \quad \Rightarrow \quad \mathscr{A}(\mathscr{K}_2) \subset \mathscr{A}(\mathscr{K}_1). \tag{2.5}$$

If $\mathscr{K}$ is a chain such that $\|[K]\| = 1$ for each $K \in J(\mathscr{K})$, then $\mathscr{A}(\mathscr{K})$ is an algebra of triangular matrices in the usual sense. Thus by (2.3), *every block-triangular matrix algebra (group) contains an algebra (group) of triangular matrices.*

Definition 2.5 may be extended as follows. Let $\mathscr{U}(I)$ denote the set of all linear subspaces of $\mathbb{R}^I$. For any subset $\mathscr{U} \subseteq \mathscr{U}(I)$ define

$$\mathscr{A}(\mathscr{U}) := \{A \in \mathscr{A}(I) \,|\, \forall U \in \mathscr{U},\ AU \subseteq U\}; \tag{2.6}$$

$\mathscr{A}(\mathscr{U})$ is again a matrix subalgebra of $\mathscr{A}(I)$, and

$$\mathscr{U}_1 \subseteq \mathscr{U}_2 \quad \Rightarrow \quad \mathscr{A}(\mathscr{U}_2) \subseteq \mathscr{A}(\mathscr{U}_1). \tag{2.7}$$

Note that for any set $\mathscr{U}$ of linear subspaces of $\mathbb{R}^I$ and any $F \in \mathscr{A}(I)^*$,

$$\mathscr{A}(F\mathscr{U}) = F\mathscr{A}(\mathscr{U})F^{-1}. \tag{2.8}$$

A set $\mathscr{U} \subseteq \mathscr{U}(I)$ is a *lattice of subspaces* if it is closed under $\cap$ and $+$ and if $\{0\}, \mathbb{R}^I \in U$.[11] A lattice $\mathscr{U}$ is called a *chain* if it is totally ordered under inclusion (hence finite).

[11] For any $\mathscr{U} \subseteq \mathscr{U}(I)$, $\mathscr{A}(\mathscr{U}) = \mathscr{A}(\mathrm{Lat}(\mathscr{U}))$, where $\mathrm{Lat}(\mathscr{U})$ is the lattice generated by $\mathscr{U}$. Thus when studying $\mathscr{A}(\mathscr{U})$, we may always assume that $\mathscr{U}$ is a lattice.

DEFINITION 2.7. Let $\mathscr{U} \subseteq \mathscr{U}(I)$ be a chain. The algebra $\mathscr{A}(\mathscr{U})$ is called the *algebra of generalized block-triangular matrices with respect to* $\mathscr{U}$. The group $\mathscr{A}(\mathscr{U})^*$ is called the *group of generalized block-triangular matrices with respect to* $\mathscr{U}$.

This definition is justified as follows. Each chain $\mathscr{K} \subseteq \mathscr{D}(I)$ determines a chain $\mathscr{U}_{\mathscr{K}} := \{U_K \mid K \in \mathscr{K}\} \subseteq \mathscr{U}(I)$ such that $\mathscr{A}(\mathscr{K}) = \mathscr{A}(\mathscr{U}_{\mathscr{K}})$. Conversely, for any chain $\mathscr{U} \subseteq \mathscr{U}(I)$ we may choose a basis of $\mathbb{R}^I$ and a chain $\mathscr{K} \subseteq \mathscr{D}(I)$ such that $\mathscr{U}$ has the form $\mathscr{U}_{\mathscr{K}}$ relative to the new basis. More precisely, there exists $F \in \mathscr{A}(I)^*$ such that $\mathscr{U} = F\mathscr{U}_{\mathscr{K}}$; hence by (2.8)

$$\mathscr{A}(\mathscr{U}) = F\mathscr{A}(\mathscr{K})F^{-1}, \tag{2.9}$$

$$\mathscr{A}(\mathscr{U})^* = F\mathscr{A}(\mathscr{K})^*F^{-1}. \tag{2.10}$$

Thus: *Every generalized block-triangular matrix algebra (group) is similar to some block-triangular matrix algebra (group).*

Suppose that $\mathscr{U}_1$ and $\mathscr{U}_2$ are chains in $\mathscr{U}(I)$ such that $\mathscr{U}_1 \subset \mathscr{U}_2$. If a chain $\mathscr{K}_2 \subseteq \mathscr{D}(I)$ and a matrix $F \in \mathscr{A}(I)^*$ are chosen such that $\mathscr{U}_2 = F\mathscr{U}_{\mathscr{K}_2}$, then there exists a subchain $\mathscr{K}_1 \subset \mathscr{K}_2$ such that $\mathscr{U}_1 = F\mathscr{U}_{\mathscr{K}_1}$. Thus by (2.5), (2.7) can be sharpened for chains as follows:

$$\mathscr{U}_1 \subset \mathscr{U}_2 \quad \Rightarrow \quad \mathscr{A}(\mathscr{U}_2) \subset \mathscr{A}(\mathscr{U}_1). \tag{2.11}$$

Since every block-triangular matrix group $\mathscr{A}(\mathscr{K})^*$ contains a group of triangular matrices, and since the latter is known (by the Cholesky decomposition) to act transitively on $\mathscr{P}(I)^*$, the following result is an immediate consequence of Remark 2.2, (2.10), and Proposition 2.3:

THEOREM 2.8. *Every generalized block-triangular group $\mathscr{A}(\mathscr{U})^*$ acts transitively on $\mathscr{P}(I)^*$.*

3. MAIN RESULTS

Our first main result is the following converse to Theorem 2.8.

THEOREM 3.1 (Existence). *Let $\mathscr{A} \subseteq \mathscr{A}(I)$ be a matrix algebra. If $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$, then there exists a chain $\mathscr{U} \subseteq \mathscr{U}(I)$ such that $\mathscr{A} = \mathscr{A}(\mathscr{U})$, i.e., $\mathscr{A}$ is a generalized block-triangular matrix algebra.*

This theorem is proved by means of Lemmas 3.2 and 3.3, whose proofs are given in Sections 4 and 5, respectively. First recall that for any subset

$\mathscr{A} \subseteq \mathscr{A}(I)$, a linear subspace $U \in \mathscr{U}(I)$ is *$\mathscr{A}$-invariant* if $\mathscr{A}U \subseteq U$. Clearly $\mathbb{R}^I$ and $\{0\}$ are $\mathscr{A}$-invariant, and U is $\mathscr{A}$-invariant iff U is $\mathrm{Alg}(\mathscr{A})$-invariant. If $\mathscr{A}$ is an algebra, then U is $\mathscr{A}$-invariant iff U is $\mathscr{A}^*$-invariant [since $\mathrm{Alg}(\mathscr{A}^*) = \mathscr{A}$].

We denote the set of all $\mathscr{A}$-invariant subspaces of $\mathbb{R}^I$ by $\mathscr{U}(\mathscr{A})$.[12] Note that $\mathscr{U}(\mathscr{A}) \subseteq \mathscr{U}(I)$ is a lattice. For any algebra $\mathscr{A} \subseteq \mathscr{A}(I)$, any lattice $\mathscr{U} \subseteq \mathscr{U}(I)$, and any $F \in \mathscr{A}(I)^*$,

$$\mathscr{U}(\mathscr{A}(\mathscr{U})) \supseteq \mathscr{U}, \tag{3.1}$$

$$\mathscr{A}(\mathscr{U}(\mathscr{A})) \supseteq \mathscr{A}, \tag{3.2}$$

$$\mathscr{A}(\mathscr{U}(\mathscr{A}(\mathscr{U}))) = \mathscr{A}(\mathscr{U}), \tag{3.3}$$

$$\mathscr{U}(\mathscr{A}(\mathscr{U}(\mathscr{A}))) = \mathscr{U}(\mathscr{A}), \tag{3.4}$$

$$F\mathscr{U}(\mathscr{A}) = \mathscr{U}(F\mathscr{A}F^{-1}). \tag{3.5}$$

Furthermore, for any two subalgebras $\mathscr{A}_1$ and $\mathscr{A}_2$ of $\mathscr{A}(I)$,

$$\mathscr{A}_1 \subseteq \mathscr{A}_2 \quad \Rightarrow \quad \mathscr{U}(\mathscr{A}_2) \subseteq \mathscr{U}(\mathscr{A}_1). \tag{3.6}$$

LEMMA 3.2. *Let $\mathscr{A} \subseteq \mathscr{A}(I)$ be a matrix algebra such that $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$. If no proper $\mathscr{A}$-invariant subspace of $\mathbb{R}^I$ exists, then $\mathscr{A} = \mathscr{A}(I)$.*

LEMMA 3.3. *Let $\mathscr{A} \subseteq \mathscr{A}(I)$ be a matrix algebra such that $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$. Suppose that U is a minimal proper $\mathscr{A}$-invariant subspace of $\mathbb{R}^I$. Then there exists a matrix $F \in \mathscr{A}(I)^*$, a proper subset $K \subset I$, and an algebra $\mathscr{B} \subseteq \mathscr{A}(K)$, such that $U = FU_K$, $\mathscr{B}^*$ acts transitively on $\mathscr{P}(K)^*$, and*

$$F^{-1}\mathscr{A}F = \left\{ \begin{pmatrix} A_{[KK]} & 0 \\ A_{[IK]} & A_{[II]} \end{pmatrix} \in \mathscr{A}(I) \,\middle|\, A_{[KK]} \in \mathscr{B}, \right.$$
$$\left. A_{[IK]} \in \mathscr{A}((I \setminus K) \times K),\ A_{[II]} \in \mathscr{A}(I \setminus K) \right\}, \tag{3.7}$$

[12] For any $\mathscr{A} \subseteq \mathscr{A}(I)$, $\mathscr{U}(\mathscr{A}) = \mathscr{U}(\mathrm{Alg}(\mathscr{A}))$, so when studying $\mathscr{U}(\mathscr{A})$ we may always assume that $\mathscr{A}$ is an algebra.

where the matrices are partitioned according to the decomposition $I = K \dot{\cup} (I \setminus K)$.

Proof of Theorem 3.1. The proof proceeds by induction on $|I| \equiv \dim \mathbb{R}^I$. If $|I| = 1$ the result is trivial. Now assume that Theorem 3.1 holds whenever $1 \leqslant |I| \leqslant p$, and consider the case $|I| = p + 1$ ($\geqslant 2$). Since $\mathbb{R}^I$ is a nonzero $\mathscr{A}$-invariant subspace and $\dim \mathbb{R}^I < \infty$, there exists a minimal nonzero $\mathscr{A}$-invariant subspace $U \subseteq \mathbb{R}^I$. If $U = \mathbb{R}^I$, then there exists no proper $\mathscr{A}$-invariant subspace, so the result follows from Lemma 3.2.

If $U \subset \mathbb{R}^I$, then by Lemma 3.3 there exists a matrix $F \in \mathscr{A}(I)^*$, a proper subset $\varnothing \subset K \subset I$, and a subalgebra $\mathscr{B} \subseteq \mathscr{A}(K)$ such that $\mathscr{B}^*$ acts transitively on $\mathscr{P}(K)^*$ and $F^{-1}\mathscr{A}F$ has the form (3.7). Since $|K| \leqslant p$, it follows from the induction hypothesis and (2.9) that there exists a chain $\mathscr{L}$ of subsets of K and a matrix $G \in \mathscr{A}(K)^*$ such that $\mathscr{B} = G\mathscr{A}(\mathscr{L})G^{-1}$. If we set

$$E = F\begin{pmatrix} G & 0 \\ 0 & 1_{I \setminus K} \end{pmatrix} \in \mathscr{A}(I)^*,$$

then from (3.7),

$$E^{-1}\mathscr{A}E = \left\{ \begin{pmatrix} A_{[KK]} & 0 \\ A_{[IK]}G & A_{[II]} \end{pmatrix} \in \mathscr{A}(I) \middle| A_{[KK]} \in \mathscr{A}(\mathscr{L}), \right.$$

$$\left. A_{[IK]} \in \mathscr{A}((I \setminus K) \times K),\ A_{[II]} \in \mathscr{A}(I \setminus K) \right\}$$

$$= \left\{ \begin{pmatrix} A_{[KK]} & 0 \\ A_{[IK]} & A_{[II]} \end{pmatrix} \in \mathscr{A}(I) \middle| A_{[KK]} \in \mathscr{A}(\mathscr{L}), \right.$$

$$\left. A_{[IK]} \in \mathscr{A}((I \setminus K) \times K),\ A_{[II]} \in \mathscr{A}(I \setminus K) \right\}$$

$$= \mathscr{A}(\mathscr{L} \cup \{I\}) \qquad [\text{by Proposition 2.4(iii)}].$$

Since $\mathscr{L} \cup \{I\}$ is a chain of subsets of I, $\mathscr{A}$ is therefore a generalized block-triangular matrix algebra, so the proof of Theorem 3.1 is complete. ■

The uniqueness of the chain $\mathscr{U}$ in Theorem 3.1 will now be established.

LEMMA 3.4. *Let $\mathscr{A} \subseteq \mathscr{A}(I)$ be a matrix algebra such that $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$. Then $\mathscr{U}(\mathscr{A})$ is a chain.*

Proof. It suffices to show that if U and V are two proper $\mathscr{A}$-invariant subspaces of $\mathbb{R}^I$ such that $U \neq V$, then either $U \cap V = U$ or $U \cap V = V$. Suppose that $U \cap V \neq U$ and $U \cap V \neq V$. Then $\mathscr{U}(\mathscr{A}) \supseteq \mathscr{U} := \{\{0\}, T, U, V, W, \mathbb{R}^I\}$, where $T := U \cap V \supseteq \{0\}$ and $W := U + V \subseteq \mathbb{R}^I$; note that $\mathscr{U}$ is a lattice but not a chain. As above (2.9), there exists $F \in \mathscr{A}(I)^*$ such that $\mathscr{U} = F\mathscr{U}_{\mathscr{K}}$, where $\mathscr{K} \subseteq \mathscr{D}(I)$ is a ring of the form $\{\varnothing, J, K, L, M, I\}$ with $K \cap L \neq K$, $K \cap L \neq L$, $J := K \cap L \supseteq \varnothing$, and $M := K \cup L \subseteq I$; again, $\mathscr{K}$ is not a chain. Since $J(\mathscr{K}) = \{J, K, L, I\}$, Proposition 2.4(iii) implies that $\mathscr{A}(\mathscr{K})$ consists of all matrices of the form

$$A = \begin{pmatrix} A_{[JJ]} & 0 & 0 & 0 \\ A_{[KJ]} & A_{[KK]} & 0 & 0 \\ A_{[LJ]} & 0 & A_{[LL]} & 0 \\ A_{[IJ]} & A_{[IK]} & A_{[IL]} & A_{[II]} \end{pmatrix},$$

where A is partitioned according to the decomposition[13] $I = [J] \dot{\cup} [K] \dot{\cup} [L] \dot{\cup} [I]$ [cf. (2.4)]. It may be shown[14] from (3.8) that $\mathscr{A}(\mathscr{K})^*$ does *not* act transitively on $\mathscr{P}(I)^*$. By (3.2), (2.7), and (2.8), however, $\mathscr{A} \subseteq \mathscr{A}(\mathscr{U}(\mathscr{A})) \subseteq \mathscr{A}(\mathscr{U}) = F\mathscr{A}(\mathscr{K})F^{-1}$; hence $\mathscr{A}(\mathscr{K})^*$ must act transitively on $\mathscr{P}(I)^*$ (by Remark 2.2 and Proposition 2.3). This contradiction establishes the result. ■

LEMMA 3.5. *Let $\mathscr{U} \subseteq \mathscr{U}(I)$ be a lattice. Then $\mathscr{U}$ is a chain iff $\mathscr{A}(\mathscr{U})^*$ acts transitively on $\mathscr{P}(I)^*$. In this case, $\mathscr{U}(\mathscr{A}(\mathscr{U})) = \mathscr{U}$.*

Proof. If $\mathscr{U}$ is a chain, then $\mathscr{A}(\mathscr{U})^*$ acts transitively on $\mathscr{P}(I)^*$ by Theorem 2.8. Conversely, if $\mathscr{A}(\mathscr{U})^*$ acts transitively on $\mathscr{P}(I)^*$, then $\mathscr{U}(\mathscr{A}(\mathscr{U}))$ is a chain by Lemma 3.4. But $\mathscr{U}(\mathscr{A}(\mathscr{U})) \supseteq \mathscr{U}$ by (3.1); hence $\mathscr{U}$ is a chain. If $\mathscr{U}(\mathscr{A}(\mathscr{U})) \supset \mathscr{U}$, then $\mathscr{A}(\mathscr{U}(\mathscr{A}(\mathscr{U}))) \subset \mathscr{A}(\mathscr{U})$ by (2.11), contradicting (3.3); thus $\mathscr{U}(\mathscr{A}(\mathscr{U})) = \mathscr{U}$. ■

[13] Note that $[J]$ and/or $[I]$ may be empty. If $[J] = \varnothing$ then $J = \varnothing$, so J does not occur in $J(\mathscr{K})$; if $[I] = \varnothing$ then $I = L$, so I does not occur (separately) in $J(\mathscr{K})$.

[14] Suppose that $\Sigma = AA'$ for some $A \in \mathscr{A}(\mathscr{K})^*$, so A has the form (3.8). Then it may be shown that $(\Sigma_M^{-1})_{[KL]} = 0$, so $\mathscr{A}(\mathscr{K})^*$ does not satisfy condition (ii) of Proposition 2.3. Alternatively, apply Remark 2.4 of Andersson and Perlman (1993).

LEMMA 3.6. *Let* $\mathscr{U} \subseteq \mathscr{U}(I)$ *be a chain and* $\mathscr{A} \subseteq \mathscr{A}(I)$ *an algebra.*

(i) $\mathscr{A}(\mathscr{U}) \subseteq \mathscr{A} \Rightarrow \mathscr{U}(\mathscr{A}) \subseteq \mathscr{U}$.
(ii) $\mathscr{A}(\mathscr{U}) \subset \mathscr{A} \Rightarrow \mathscr{U}(\mathscr{A}) \subset \mathscr{U}$.

Proof. (i): By (3.6) and Lemma 3.5, $\mathscr{A}(\mathscr{U}) \subseteq \mathscr{A} \Rightarrow \mathscr{U}(\mathscr{A}) \subseteq \mathscr{U}(\mathscr{A}(\mathscr{U})) = \mathscr{U}$.

(ii): By (i), $\mathscr{U}(\mathscr{A}) \subseteq \mathscr{U}$. If $\mathscr{U}(\mathscr{A}) = \mathscr{U}$, then by (3.2), $\mathscr{A} \subseteq \mathscr{A}(\mathscr{U}(\mathscr{A})) = \mathscr{A}(\mathscr{U})$, a contradiction; hence $\mathscr{U}(\mathscr{A}) \subset \mathscr{U}$.

LEMMA 3.7. *When* $\mathscr{A}_1$ *and* $\mathscr{A}_2$ *are subalgebras of* $\mathscr{A}(I)$ *such that* $\mathscr{A}_1^*$ *acts transitively on* $\mathscr{P}(I)^*$, (3.6) *may be sharpened as follows*:

$$\mathscr{A}_1 \subset \mathscr{A}_2 \quad \Rightarrow \quad \mathscr{U}(\mathscr{A}_2) \subset \mathscr{U}(\mathscr{A}_1). \tag{3.9}$$

Proof. By Theorem 3.1, there exists a chain $\mathscr{U}$ such that $\mathscr{A}_1 = \mathscr{A}(\mathscr{U})$. By Lemma 3.6(ii) and (3.1), $\mathscr{U}(\mathscr{A}_2) \subset \mathscr{U} \subseteq \mathscr{U}(\mathscr{A}(\mathscr{U})) = \mathscr{U}(\mathscr{A}_1)$.

THEOREM 3.8 (Uniqueness). *Let* $\mathscr{A} \subseteq \mathscr{A}(I)$ *be a matrix algebra such that* $\mathscr{A}^*$ *acts transitively on* $\mathscr{P}(I)^*$. *Then* $\mathscr{U}(\mathscr{A})$ *is a chain,* $\mathscr{A}(\mathscr{U}(\mathscr{A})) = \mathscr{A}$, *and* $\mathscr{U}(\mathscr{A})$ *is the unique lattice* $\mathscr{V} \subseteq \mathscr{U}(I)$ *such that* $\mathscr{A}(\mathscr{V}) = \mathscr{A}$.

Proof. By Lemma 3.4, $\mathscr{U}(\mathscr{A})$ is a chain. By (3.2), $\mathscr{A}(\mathscr{U}(\mathscr{A})) \supseteq \mathscr{A}$. If $\mathscr{A}(\mathscr{U}(\mathscr{A})) \supset \mathscr{A}$ then $\mathscr{U}(\mathscr{A}(\mathscr{U}(\mathscr{A}))) \subset \mathscr{U}(\mathscr{A})$ by Lemma 3.7, which contradicts (3.3); hence $\mathscr{A}(\mathscr{U}(\mathscr{A})) = \mathscr{A}$. If $\mathscr{V} \subseteq \mathscr{U}(I)$ is a lattice such that $\mathscr{A}(\mathscr{V}) = \mathscr{A}$ then $\mathscr{U}(\mathscr{A}) = \mathscr{U}(\mathscr{A}(\mathscr{V})) \supseteq \mathscr{V}$ by (3.1); hence $\mathscr{V}$ is a chain. Therefore $\mathscr{U}(\mathscr{A}(\mathscr{V})) = \mathscr{V}$ by Lemma 3.5, so $\mathscr{U}(\mathscr{A}) = \mathscr{V}$. ■

The following two corollaries have statistical applications (cf. Andersson, Marden, and Perlman, 1994).

COROLLARY 3.9. *Let* $\mathscr{U} \subseteq \mathscr{U}(I)$ *be a chain and* $\mathscr{A} \subseteq \mathscr{A}(I)$ *an algebra. Then* $\mathscr{A} \supseteq \mathscr{A}(\mathscr{U})$ *if and only if* $\mathscr{A} = \mathscr{A}(\mathscr{V})$ *for some subchain* $\mathscr{V} \subseteq \mathscr{U}$.

Proof. "If": apply (2.7).

"Only if": by Theorem 2.8 and Remark 2.2, $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$. Now set $\mathscr{V} = \mathscr{U}(\mathscr{A})$, so the result follows from Theorem 3.8 and Lemma 3.6(i). ■

COROLLARY 3.10.

(i) *Let* $\mathscr{U} \subseteq \mathscr{U}(I)$ *be a chain and* $\mathscr{A} \subseteq \mathscr{A}(I)$ *an algebra such that* $\mathscr{A} \subseteq \mathscr{A}(\mathscr{U})$. *Then* $\mathscr{A}^*$ *acts transitively on* $\mathscr{P}(I)^*$ *if and only if* $\mathscr{A} = \mathscr{A}(\mathscr{V})$ *for some chain* $\mathscr{V} \supseteq \mathscr{U}$.

(ii) *Let* $\mathcal{K}$ *be a chain of subsets of* I, *and* $\mathcal{A} \subseteq \mathcal{A}(I)$ *an algebra, such that* $\mathcal{A} \subseteq \mathcal{A}(\mathcal{K})$. *Then* $\mathcal{A}^*$ *acts transitively on* $\mathcal{P}(I)^*$ *if and only if* $\mathcal{A} = F\mathcal{A}(\mathcal{L})F^{-1}$ *for some chain* $\mathcal{L} \supseteq \mathcal{K}$ *and some* $F \in \mathcal{A}(\mathcal{K})^*$.

Proof. (i), "if": Apply Theorem 2.8. "Only if": Set $\mathcal{V} = \mathcal{U}(\mathcal{A})$, a chain by Lemma 3.4. Then $\mathcal{A} = \mathcal{A}(\mathcal{V})$ by Lemma 3.5, and $\mathcal{V} \supseteq \mathcal{U}$ by (3.6).

(ii): Apply (i) with $\mathcal{U} = \mathcal{U}_{\mathcal{K}}$, and note that $\mathcal{V} \supseteq \mathcal{U}_{\mathcal{K}}$ if and only if $\mathcal{V} = F\mathcal{U}_{\mathcal{L}}$ for some chain $\mathcal{L} \supseteq \mathcal{K}$ and some $F \in \mathcal{A}(\mathcal{K})^*$. Now apply (2.8). ∎

We conclude this section with several related results of possible interest.

LEMMA 3.11. *Let* $\mathcal{U} \subseteq \mathcal{U}(I)$ *be a lattice and* $\mathcal{A} \subseteq \mathcal{A}(I)$ *an algebra such that* $\mathcal{A}^*$ *acts transitively on* $\mathcal{P}(I)^*$.

(i) $\mathcal{U} \subseteq \mathcal{U}(\mathcal{A}) \Leftrightarrow \mathcal{A}(\mathcal{U}) \supseteq \mathcal{A}$.
(ii) $\mathcal{U} \subset \mathcal{U}(\mathcal{A}) \Leftrightarrow \mathcal{A}(\mathcal{U}) \supset \mathcal{A}$.

Proof. (i), $\Rightarrow$: Apply (2.7) and (3.2). $\Leftarrow$: By (3.6), $\mathcal{U}(\mathcal{A}(\mathcal{U})) \subseteq \mathcal{U}(\mathcal{A})$. But by Remark 2.2, $\mathcal{A}(\mathcal{U})^*$ acts transitively on $\mathcal{P}(I)^*$; hence $\mathcal{U} = \mathcal{U}(\mathcal{A}(\mathcal{U}))$ by Lemma 3.5.

(ii), $\Rightarrow$: by (i), $\mathcal{A}(\mathcal{U}) \supseteq \mathcal{A}$. If $\mathcal{A}(\mathcal{U}) = \mathcal{A}$ then $\mathcal{U} = \mathcal{U}(\mathcal{A})$ by Theorem 3.8, which contradicts the hypothesis; hence $\mathcal{A}(\mathcal{U}) \supset \mathcal{A}$. $\Leftarrow$: By (i), $\mathcal{U} \subseteq \mathcal{U}(\mathcal{A})$. If $\mathcal{U} = \mathcal{U}(\mathcal{A})$ then $\mathcal{A}(\mathcal{U}) = \mathcal{A}(\mathcal{U}(\mathcal{A})) = \mathcal{A}$ by Theorem 3.8, a contradiction. ∎

COROLLARY 3.12. *Let* $\mathcal{U} \subseteq \mathcal{U}(I)$ *be a lattice and* $\mathcal{A} \subseteq \mathcal{A}(I)$ *an algebra such that* $\mathcal{A}^*$ *acts transitively on* $\mathcal{P}(I)^*$. *Then* $\mathcal{U} \subseteq \mathcal{U}(\mathcal{A})$ *if and only if* $\mathcal{U} = \mathcal{U}(\mathcal{B})$ *for some algebra* $\mathcal{B} \supseteq \mathcal{A}$.

Proof. "If": Use (2.7).

"Only if": by Lemma 3.4, $\mathcal{U}(\mathcal{A})$ is a chain; hence $\mathcal{U}$ is a chain. Now set $\mathcal{B} = \mathcal{A}(\mathcal{U})$, and apply Lemma 3.5 and Lemma 3.11.

COROLLARY 3.13. *Let* $\mathcal{U} \subseteq \mathcal{U}(I)$ *be a lattice and* $\mathcal{A} \subseteq \mathcal{A}(I)$ *an algebra such that* $\mathcal{A}^*$ *acts transitively on* $\mathcal{P}(I)^*$ *and* $\mathcal{U} \supseteq \mathcal{U}(\mathcal{A})$. *Then* $\mathcal{U}$ *is a chain if and only if* $\mathcal{U} = \mathcal{U}(\mathcal{B})$ *for some algebra* $\mathcal{B} \subseteq \mathcal{A}$ *such that* $\mathcal{B}^*$ *acts transitively on* $\mathcal{P}(I)^*$.

Proof. "If": Apply Theorem 3.8 to $\mathcal{B}$.

"Only if": Set $\mathcal{B} = \mathcal{A}(\mathcal{U})$. Then $\mathcal{B}^*$ acts transitively on $\mathcal{P}(I)^*$ by Theorem 2.8, $\mathcal{U} = \mathcal{U}(\mathcal{B})$ by Lemma 3.5, and $\mathcal{B} \subseteq \mathcal{A}$ by Theorem 3.8. ∎

4. PROOF OF LEMMA 3.2

Lemma 3.2 is an immediate consequence of the following result:

LEMMA 4.1 (Burnside's theorem for real matrices). *Let $\mathscr{A} \subseteq \mathscr{A}(I)$ be a matrix algebra such that no proper $\mathscr{A}$-invariant subspace of $\mathbb{R}^I$ exists. Then either $\mathscr{A} = \mathscr{A}(I)$, or else there exists $F \in \mathscr{A}(I)^*$ such that $\mathscr{A} = F^{-1}\mathscr{A}_{\mathbb{C}}(I)F$ or $\mathscr{A} = F^{-1}\mathscr{A}_{\mathbb{H}}(I)F$, where*

$$\mathscr{A}_{\mathbb{C}}(I) = \left\{ \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \in \mathscr{A}(I) \middle| A, B \in \mathscr{A}(N) \right\}, \qquad I = N \dot{\cup} N, \quad (4.1)$$

$$\mathscr{A}_{\mathbb{H}}(I) = \left\{ \begin{pmatrix} A & -B & -C & -D \\ B & A & D & -C \\ C & -D & A & B \\ D & C & -B & A \end{pmatrix} \in \mathscr{A}(I) \middle| A, B, C, D \in \mathscr{A}(Q) \right\},$$

$$I = Q \dot{\cup} Q \dot{\cup} Q \dot{\cup} Q. \quad (4.2)$$

Proof. The commutant algebra $\mathscr{B} \equiv \mathscr{B}(\mathscr{A})$ and bicommutant algebra $\mathscr{C} \equiv \mathscr{C}(\mathscr{A})$ are defined, respectively, as

$$\mathscr{C} = \{ C \in \mathscr{A}(I) \,|\, CA = AC \ \forall A \in \mathscr{A} \},$$

$$\mathscr{B} = \{ B \in \mathscr{A}(I) \,|\, BC = CB \ \forall C \in \mathscr{C} \}.$$

By Schur's lemma, $\mathscr{C}$ is a division algebra over $\mathbb{R}$; hence by Frobenius' theorem, $\mathscr{C}$ is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$. If $\mathscr{C} \cong \mathbb{R}$ then $\mathscr{C} = \{\alpha 1_I | \alpha \in \mathbb{R}\}$, so $\mathscr{B} = \mathscr{A}(I)$.

If $\mathscr{C} \cong \mathbb{C}$ then $\mathscr{C} = \{\alpha 1_I + \beta J | \alpha, \beta \in \mathbb{R}\}$, where $J \in \mathscr{A}(I)^*$ satisfies $J^2 = -1_I$, and $\mathbb{R}^I$ may be considered as a vector space over $\mathscr{C}$. Let $(e_n | n \in N)$ be a basis for $\mathbb{R}^I$ over $\mathscr{C}$. Then $((e_n | n \in N), (Je_n | n \in N))$ is a basis for $\mathbb{R}^I$ over $\mathscr{R} := \{\alpha 1_I | \alpha \in \mathbb{R}\} \cong \mathbb{R}$, so $I = N \dot{\cup} N$. Now choose $F \in \mathscr{A}(I)^*$ such that $((Fe_n | n \in N), (FJe_n | n \in N))$ is the canonical basis for $\mathbb{R}^I$ over $\mathbb{R}$. Then it is easily seen that

$$J = F^{-1} \begin{pmatrix} 0 & -1_N \\ 1_N & 0 \end{pmatrix} F; \quad (4.3)$$

hence $\mathscr{B} = F^{-1}\mathscr{A}_{\mathbb{C}}(I)F$.

If $\mathscr{E} \cong \mathbb{H}$ then $\mathscr{E} = \{\alpha 1_I + \beta J + \gamma K + \delta L | \alpha, \beta, \gamma, \delta \in \mathbb{R}\}$, where $J, K, L \in \mathscr{A}(I)^*$ satisfy $J^2 = K^2 = L^2 = -1_I$, $JK = -KJ = L$, $KL = -LK = J$, $LJ = -LJ = K$, and $\mathbb{R}^I$ may be considered as a vector space over $\mathscr{E}$. Let $(e_q | q \in Q)$ be a basis for $\mathbb{R}^I$ over $\mathscr{E}$. Then $((e_q | q \in Q), (Je_q | q \in Q), (Ke_q | q \in Q), (Le_q | q \in Q))$ is a basis for $\mathbb{R}^I$ over $\mathscr{R}$, so $I = Q \dot{\cup} Q \dot{\cup} Q \dot{\cup} Q$. Now choose $F \in \mathscr{A}(I)^*$ such that $((Fe_q | q \in Q), (FJe_q | q \in Q), (FKe_q | q \in Q), (FLe_q | q \in Q))$ is the canonical basis for $\mathbb{R}^I$ over $\mathbb{R}$. Then

$$J = F^{-1} \begin{pmatrix} 0 & -1_Q & 0 & 0 \\ 1_Q & 0 & 0 & 0 \\ 0 & 0 & 0 & -1_Q \\ 0 & 0 & 1_Q & 0 \end{pmatrix} F, \tag{4.4}$$

$$K = F^{-1} \begin{pmatrix} 0 & 0 & -1_Q & 0 \\ 0 & 0 & 0 & 1_Q \\ 1_Q & 0 & 0 & 0 \\ 0 & -1_Q & 0 & 0 \end{pmatrix} F, \tag{4.5}$$

$$L = F^{-1} \begin{pmatrix} 0 & 0 & 0 & -1_Q \\ 0 & 0 & -1_Q & 0 \\ 0 & 1_Q & 0 & 0 \\ 1_Q & 0 & 0 & 0 \end{pmatrix} F; \tag{4.6}$$

hence $\mathscr{B} = F^{-1} \mathscr{A}_{\mathbb{H}}(I) F$.

By the bicommutant theorem (cf. Bourbaki, 1958, §4, No. 2, Corollary 1), however, $\mathscr{B} = \mathscr{A}$. This completes the proof. ■

Proof of Lemma 3.2. It is easily verified that $[\mathscr{A}_{\mathbb{C}}(I)][\mathscr{A}_{\mathbb{C}}(I)]^t \subseteq \mathscr{A}_{\mathbb{C}}(I) \cap \mathscr{P}(I) \subset \mathscr{P}(I)$, so by Proposition 2.3, $\mathscr{A}_{\mathbb{C}}(I)^*$ does not act transitively on $\mathscr{P}(I)^*$. Similarly, $[A_{\mathbb{H}}(I)][\mathscr{A}_{\mathbb{H}}(I)]^t \subseteq \mathscr{A}_{\mathbb{H}}(I) \cap \mathscr{P}(I) \subset \mathscr{P}(I)$, so $\mathscr{A}_{\mathbb{H}}(I)^*$ does not act transitively on $\mathscr{P}(I)^*$. Thus, under the hypothesis of Lemma 3.2, Lemma 4.1 implies that $\mathscr{A} = \mathscr{A}(I)$. The proof is complete. ■

REMARK 4.2. If $|I|$ is odd, then Lemma 3.2 is true without the assumption that $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$, since in that case it is not possible that $\mathscr{A} = F^{-1}\mathscr{A}_{\mathbb{C}}(I)F$ or $\mathscr{A} = F^{-1}\mathscr{A}_{\mathbb{H}}(I)F$. For an algebra $\mathscr{A}$ of matrices with complex entries, Burnside's theorem[15] states that if no proper $\mathscr{A}$-invariant subspace of $\mathbb{C}^I$ exists, then $\mathscr{A} = \mathscr{A}(I)$. Thus in the complex case, Lemma 3.2 is true for *every* I, without the assumption that $\mathscr{A}^*$ acts transitively on $\mathscr{P}(I)^*$.

5. PROOF OF LEMMA 3.3

Without loss of generality we may assume that $U = U_K$ for some proper subset $K \subset I$. Then $\mathscr{A} \subseteq \mathscr{A}(\mathscr{K})$, where $\mathscr{K} := \{\varnothing, K, I\}$ and $\mathscr{A}(\mathscr{K})$ are as in Example 2.6. We shall show that (3.7) holds with $F = 1_I$. For any subset $N \subset I$ let $O(N)$ denote the group of all real orthogonal $N \times N$ matrices, and let 0_N denote the $N \times N$ zero matrix.

FACT 1. $\operatorname{Diag}(1_K, 0_{I\setminus K}) \in \mathscr{A}$, $\operatorname{Diag}(0_K, 1_{I\setminus K}) \in \mathscr{A}$.

Proof. Since $\operatorname{Diag}(1_K, 0_{I\setminus K}) \in \mathscr{P}(I)$, the transitivity of $\mathscr{A}$ and Proposition 2.3(i) imply that $\exists A \in \mathscr{A} \ni AA^t = \operatorname{Diag}(1_K, 0_{I\setminus K})$. As $\mathscr{A} \subseteq \mathscr{A}(\mathscr{K})$, this implies that $A = \operatorname{Diag}(\Gamma, 0_{I\setminus K})$ for some $\Gamma \in O(K)$. Similarly, since $\operatorname{Diag}(\varepsilon 1_K, 1_{I\setminus K}) \in \mathscr{P}(I)$ $\forall \varepsilon > 0$, $\exists \Gamma_\varepsilon \in O(K)$, $\psi_\varepsilon \in O(I\setminus K)$ such that $\operatorname{Diag}(\sqrt{\varepsilon}\,\Gamma_\varepsilon, \psi_\varepsilon) \in \mathscr{A}$. Because $O(K)$ and $O(I\setminus K)$ are compact and $\mathscr{A}$ is closed in $\mathscr{A}(I)$, $\exists \psi \in O(I\setminus K)$ such that $\operatorname{Diag}(0_K, \psi) \in \mathscr{A}$. Since $\mathscr{A}$ is an algebra, $\operatorname{Diag}(\Gamma, \psi) \in \mathscr{A}$; hence $\operatorname{Diag}(\Gamma^t, \psi^t) \equiv \operatorname{Diag}(\Gamma, \psi)^{-1} \in \mathscr{A}$. Thus

$$\operatorname{Diag}(1_K, 0_{I\setminus K}) \equiv \operatorname{Diag}(\Gamma, 0_{I\setminus K}) \operatorname{Diag}(\Gamma^t, \psi^t) \in \mathscr{A},$$

$$\operatorname{Diag}(0_K, 1_{I\setminus K}) \equiv \operatorname{Diag}(0_K, \psi) \operatorname{Diag}(\Gamma^t, \psi^t) \in \mathscr{A}. \qquad \blacksquare$$

FACT 2. *Define* $\mathscr{A}_{[KK]} = \{A_{[KK]} | A \in \mathscr{A}\}$, $\mathscr{A}_{[IK]} = \{A_{[IK]} | A \in \mathscr{A}\}$, $\mathscr{A}_{[II]} = \{A_{[II]} | A \in \mathscr{A}\}$. *Then*

$$\mathscr{A}_{[KK]} = \tilde{\mathscr{A}}_{[KK]} := \left\{A_{[KK]} \in \mathscr{A}(K) \middle| \operatorname{Diag}(A_{[KK]}, 0_{I\setminus K}) \in \mathscr{A}\right\},$$

$$\mathscr{A}_{[IK]} = \tilde{\mathscr{A}}_{[IK]} := \left\{A_{[IK]} \in \mathscr{A}((I\setminus K) \times K) \middle| \begin{pmatrix} 0 & 0 \\ A_{[IK]} & 0 \end{pmatrix} \in \mathscr{A}\right\},$$

$$\mathscr{A}_{[II]} = \tilde{\mathscr{A}}_{[II]} := \left\{A_{[II]} \in \mathscr{A}(I\setminus K) \middle| \operatorname{Diag}(0_K, A_{[II]}) \in \mathscr{A}\right\},$$

[15] For example, see Jacobson (1953), Rosenthal (1984), or Gohberg, Lancaster, and Rodman (1986).

and

$$\mathscr{A} = \left\{ \begin{pmatrix} A_{[KK]} & 0 \\ A_{[IK]} & A_{[II]} \end{pmatrix} \in \mathscr{A}(I) \,\middle|\, A_{[KK]} \in \mathscr{A}_{[KK]}, \right.$$

$$\left. A_{[IK]} \in \mathscr{A}_{[IK]},\ A_{[II]} \in \mathscr{A}_{[II]} \right\}. \tag{5.1}$$

Also, $\mathscr{A}_{[KK]}$ *and* $\mathscr{A}_{[II]}$ *are matrix algebras.*

Proof. Clearly $\mathscr{A}_{[KK]} \supseteq \tilde{\mathscr{A}}_{[KK]}$. Conversely, if $A \in \mathscr{A}$, then by Fact 1,

$$\mathrm{Diag}(A_{[KK]}, 0_{I \setminus K}) \equiv \mathrm{Diag}(1_K, 0_{I \setminus K})\ A\ \mathrm{Diag}(1_K, 0_{I \setminus K}) \in \mathscr{A},$$

so $A_{[KK]} \in \tilde{\mathscr{A}}_{[KK]}$; hence $\mathscr{A}_{[KK]} = \tilde{\mathscr{A}}_{[KK]}$. Similarly, $\mathscr{A}_{[IK]} = \tilde{\mathscr{A}}_{[IK]}$ and $\mathscr{A}_{[II]} = \tilde{\mathscr{A}}_{[II]}$. Then (5.1) follows from these identities and the fact that $\mathscr{A}$ is closed under addition. Since $\tilde{\mathscr{A}}_{[KK]}$ and $\tilde{\mathscr{A}}_{[II]}$ are matrix algebras,[16] the final assertion is immediate. ■

FACT 3. $\{(A_{[KK]}A^t_{[KK]}, A_{[IK]}A^{-1}_{[KK]}, A_{[II]}A^t_{[II]}) | A \in \mathscr{A}^*\} = \mathscr{P}(K)^* \times \mathscr{M}((I \setminus K) \times K) \times \mathscr{P}(I \setminus K)^*$. *Therefore,* $(\mathscr{A}_{[KK]})^*$ *acts transitively on* $\mathscr{P}(K)^*$, *and* $(\mathscr{A}_{[II]})^*$ *acts transitively on* $\mathscr{P}(I \setminus K)^*$. *Also, no proper* $\mathscr{A}_{[II]}$*-invariant subspace of* $\mathbb{R}^{I \setminus K}$ *exists, so* $\mathscr{A}_{[II]} = \mathscr{A}(I \setminus K)$.

Proof. The first assertion follows from the transitivity of $\mathscr{A}^*$, i.e., $\mathscr{A}^*(\mathscr{A}^*)^t = \mathscr{P}(I)^*$, and from the well-known result[17] that the mapping

$$\begin{gathered} \mathscr{P}(I)^* \to \mathscr{P}(K)^* \times \mathscr{M}((I \setminus K) \times K) \times \mathscr{P}(I \setminus K)^*, \\ \Sigma \to \left(\Sigma_{[KK]}, \Sigma_{[IK]}\Sigma^{-1}_{[KK]}, \Sigma_{[II]} - \Sigma_{[IK]}\Sigma^{-1}_{[KK]}\Sigma_{[KI]}\right) \end{gathered} \tag{5.2}$$

is a bijection. The second assertion follows from the first on considering the subset $\mathscr{P}(K)^* \times \{0_{[IK]}\} \times \mathscr{P}(I \setminus K)^*$. Thirdly, if V is a proper $\mathscr{A}_{[II]}$-invariant subspace of $\mathbb{R}^{I \setminus K}$, then $\{0\} \times V$ is a proper $\mathscr{A}$-invariant subspace of $\{0\} \times \mathbb{R}^{I \setminus K} \equiv U_K$, contradicting the minimality of $U \equiv U_K$ (here 0 denotes the zero vector in $\mathbb{R}^k$). Thus $\mathscr{A}_{[II]} = \mathscr{A}(I \setminus K)$ by Lemma 3.2. ■

[16] By Fact 1, $1_K \in \tilde{\mathscr{A}}_{[KK]}$ and $1_{I \setminus K} \in \tilde{\mathscr{A}}_{[II]}$.

[17] See Andersson and Perlman (1993, Theorem 2.2) for a more general result.

FACT 4. $\mathscr{A}_{[IK]} = \mathscr{A}((I \setminus K) \times K)$.

Proof. For any $B \in \mathscr{A}((I \setminus K) \times K)$ we have $(1_K, B, 1_{I \setminus K}) \in \mathscr{P}(K)^* \times \mathscr{A}((I \setminus K) \times K) \times \mathscr{P}(I \setminus K)^*$, so by Fact 3, $\exists A \in \mathscr{A}$ such that $(A_{[KK]}A^t_{[KK]}, A_{[IK]}A^{-1}_{[KK]}, A_{[II]}A^t_{[II]}) = (1_K, B, 1_{I \setminus K})$. Thus $A_{[KK]} \in O(K)$ and $A_{[II]} \in O(I \setminus K)$. Since $\mathscr{A}_{[KK]}$ is an algebra, $A^t_{[KK]} \equiv A^{-1}_{[KK]} \in \mathscr{A}_{[KK]}$ and $\mathrm{Diag}(A^t_{[KK]}, 0_{I \setminus K}) \in \mathscr{A}$. Therefore

$$\begin{pmatrix} 1_K & 0 \\ B & 0 \end{pmatrix} = A\,\mathrm{Diag}\left(A^t_{[KK]}, 0_{I \setminus K}\right) \in \mathscr{A};$$

hence by Fact 1,

$$\begin{pmatrix} 0 & 0 \\ B & 0 \end{pmatrix} = \mathrm{Diag}\left(0_K, 1_{I \setminus K}\right) \begin{pmatrix} 1_K & 0 \\ B & 0 \end{pmatrix} \in \mathscr{A}.$$

Thus $B \in \mathscr{A}_{[IK]}$ by Fact 2, which completes the proof. ∎

Proof of Lemma 3.3. If we set $F = 1_I$ and $\mathscr{B} = \mathscr{A}_{[KK]}$, then the desired result follows from Facts 2, 3, and 4. ∎

REFERENCES

Anderson, T. W. 1984. *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, New York.

Andersson, S. A., Marden, J. I., and Perlman, M. D. 1994. *Totally Ordered Multivariate Linear Models and Testing Problems*, monograph in preparation.

Andersson, S. A. and Perlman, M. D. 1993. Lattice models for conditional independence in a multivariate normal distribution, *Ann. Statist.*, 21:1318–1358.

Bourbaki, N. 1958. *Algèbre, Livre II, Modules et Anneaux Semi-simples*, Élém. de Math. XXIII, Hermann, Paris, Chapter 8.

Eaton, M. L. 1983. *Multivariate Statistics: A Vector Space Approach*, Wiley, New York.

Eaton, M. L. and Kariya, T. 1983. Multivariate tests with incomplete data, *Ann. Statist.* 11:654–665.

Giri, N. 1968. Locally and asymptotically minimax tests of a multivariate problem, *Ann. Math. Statist.* 39:171–178.

Giri, N. 1977. *Multivariate Statistical Inference*, Academic, New York.

Giri, N., Kiefer, J., and Stein, C. 1963. Minimax character of Hotelling's T^2-test in the simplest case, *Ann. Math. Statist.* 34:1524–1535.

Gohberg, I., Lancaster, P., and Rodman, L. 1986. *Invariant Subspaces of Matrices with Applications*, Wiley, New York.

Jacobson, N. 1953. *Lectures in Abstract Algebra II: Linear Algebra*, Van Nostrand, Princeton, N.J.

James, N. and Stein, C. 1961. Estimation with quadratic loss, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* (J. Neyman, Ed.), Vol. 1, pp. 361–379.

Marden, J. I. 1983. Admissibility of invariant tests in the general multivariate analysis of variance, *Ann. Statist.* 11:1086–1099.

Marden, J. I. and Perlman, M. D. 1980. Invariant tests for means with covariates, *Ann. Statist.* 8:25–63.

Marden, J. I. and Perlman, M. D. 1990. On the inadmissibility of step-down procedures for the Hotelling T^2 problem, *Ann. Statist.* 18:172–190.

Olkin I. and Selliah, J. B. 1977. Estimating covariances in a multivariate normal distribution, in *Statistical Decision Theory and Related Topics II* (S. S. Gupta, Ed.), pp. 313–326, Academic Press, New York.

Rosenthal, E. 1984. A remark on Burnside's theorem on matrix algebras, *Linear Algebra Appl.* 63:175–177.

Received 31 August 1992