Inequalities in Statistics and Probability
IMS Lecture Notes–Monograph Series Vol. 5 (1984), 26–34

INVARIANT ORDERING AND ORDER PRESERVATION[1]

BY D. R. JENSEN
Virginia Polytechnic Institute and State University

Suppose $\mathcal{G}$ is a group of one-to-one transformations of a set $\mathcal{X}$ onto $\mathcal{X}$, M is maximal invariant taking values in an ordered set $(\mathcal{M}, \succeq_M)$, and $\succeq$ is an ordering on $\mathcal{X}$ induced from $\succeq_M$. Properties of $(\mathcal{X}, \succeq)$ are studied in Part I including lattice properties and order preservation. Examples include an ordering on $\mathcal{F}_{n\times m}$ having properties of Loewner's (1934) ordering for Hermitian varieties, a unitary ordering on $\mathcal{F}_{n\times m}$ giving a lattice, and orderings on $\mathcal{R}^n$ invariant under various groups. Applications to a variety of problems in statistics and applied probability are given in Part II..

PART I. THEORY

1. Introduction. Many problems exhibit symmetries as invariance under a group $\mathcal{G}$ acting on a set $\mathcal{X}$. Invariance principles require that solutions be invariant, and reduction by invariance preserves essentials while discarding irrelevant details. Because order relations often assume a prominent role in the analysis of such problems, it is instructive to consider orderings symmetric under $\mathcal{G}$.

Often $\mathcal{X}$ is finite-dimensional; examples are the Euclidean space $\mathcal{R}^n$, the linear space $\mathcal{F}_{n\times m}$ of $(n\times m)$ matrices over the complex field $\mathcal{C}$, the Hermitian $(n\times n)$ matrices $\mathcal{H}_n$, and the cone $\mathcal{H}_n^+$ of positive semidefinite Hermitian varieties. Typical groups of transformations are the classical groups. An ordering on $\mathcal{H}_n$ in wide usage was studied by Loewner (1934) in which $\mathbf{A} \succeq_L \mathbf{B}$ on $\mathcal{H}_n$ if and only if $\mathbf{A} - \mathbf{B} \in \mathcal{H}_n^+$. This ordering is invariant under the general linear group $Gl(n)$ acting on $\mathcal{H}_n$ by congruence, for $\mathbf{A} \succeq_L \mathbf{B}$ on $\mathcal{H}_n$ if and only if $\mathbf{CAC^*} \succeq_L \mathbf{CBC^*}$ on $\mathcal{H}_n$ for every $\mathbf{C} \in Gl(n)$, with $\mathbf{C^*}$ the conjugate transpose of $\mathbf{C}$. The relation $\succeq_L$ as an ordering on $\mathcal{F}_{n\times n}$ was considered by Hartwig (1976).

Here we study symmetric orderings induced through maximal invariants, the preservation of such orderings, and the possible transitivity of lattice properties. Our principal motivation stems from needed orderings on all of $\mathcal{F}_{n\times m}$ and not just $\mathcal{F}_{n\times n}$ or its Hermitian varieties.

2. The Basic Results. A set $\mathcal{X}$ together with a binary relation $\succeq$ is said to be *linearly ordered* if the relation is reflexive, transitive, antisymmetric, and complete. The relation is a *partial ordering* if it is reflexive, transitive, and antisymmetric, and a *preordering* if it is reflexive and transitive. A partially ordered set $(\mathcal{X}, \succeq)$ is a *lower semi-lattice* if for any two elements x,y there is an element $v = x \wedge y \in \mathcal{X}$ that is a greatest lower bound for x,y; an *upper semi-lattice* if there is a least upper bound $u = x \vee y$ for x,y in $\mathcal{X}$; and a *lattice* if it is both a lower and upper semi-lattice.

Let $\mathcal{G}$ be a group of one-to-one transformations from $\mathcal{X}$ onto $\mathcal{X}$. A function f on $\mathcal{X}$ is said to be *invariant* under $\mathcal{G}$ if, for any $(x,g) \in \mathcal{X} \times \mathcal{G}$, $f(gx) = f(x)$, and to be *maximal invariant* if it is invariant and if $f(x) = f(y)$ implies $y = gx$ for some $g \in \mathcal{G}$. The $\mathcal{G}$*-orbit* of $x_0 \in \mathcal{X}$

[1] This work was supported in part by the U.S. Army Research Office through Grant No. DAAG–29–78–G–0172.

AMS 1980 subject classifications. Primary, 06A10; Secondary, 15A45.

Key words and phrases: Invariant orderings, maximal invariants, monotone functions, orderings on $\mathcal{F}_{n\times m}$, Moore-Penrose inverses, applications.

is the set $O(x_0) = \{x \in \mathcal{X} | x = gx_0,\ g \in \mathcal{G}\}$ in $\mathcal{X}$. An invariant function is constant on each orbit; a maximal invariant function takes distinct values on different orbits. The orbits are equivalence classes under the relation $y \equiv x \bmod \mathcal{G}$, in one-to-one correspondence with the values of a maximal invariant. If the image $\mathcal{V}$ of $\mathcal{X}$ under f has a partial ordering $\succeq_V$, the mapping f is said to be *monotone* if f is order-preserving, i.e. if $x \succeq y$ on $\mathcal{X}$ implies $f(x) \succeq_V f(y)$ on $\mathcal{V}$. In particular, a real function ϕ is monotone on $(\mathcal{X}, \succeq)$ if $x \succeq y$ implies $\phi(x) \geq \phi(y)$; if ϕ is a norm it is a *monotone norm*.

Henceforth $M(x)$ on $\mathcal{X}$ is a maximal invariant function with range $\mathcal{M}$. If in addition $(\mathcal{M}, \succeq_M)$ is ordered, then a binary relation on $\mathcal{X}$ may be induced as follows.

Definition 1. Let $M: \mathcal{X} \to \mathcal{M}$ be maximal invariant under $\mathcal{G}$. If $(\mathcal{M}, \succeq_M)$ is ordered, then x,y are said to be related as $x \succeq y$ on $\mathcal{X}$ if and only if $M(x) \succeq_M M(y)$ on $\mathcal{M}$.

The following theorem is basic. It asserts that the induced relation on $\mathcal{X}$ is an ordering, that this ordering is invariant, that $(\mathcal{X}, \succeq)$ may inherit lattice properties from $(\mathcal{M}, \succeq_M)$, and that the functions monotone on $(\mathcal{X}, \succeq)$ may be characterized. That a property holds up to equivalence means that the conventional definition applies when elements on the same $\mathcal{G}$-orbit are identified.

THEOREM 1. *Let $\succeq$ be a binary relation on $\mathcal{X}$ induced as in Definition 1.* (i) ***The relation is invariant in the sense that $x \succeq y$ on $\mathcal{X}$ if an only if $gx \succeq g'y$ for any $g,g' \in \mathcal{G}$.*** (ii) ***If $(\mathcal{M}, \succeq_M)$ is partially ordered, then $(\mathcal{X}, \succeq)$ is preordered and is antisymmetric up to equivalence.*** (iii) ***If $(\mathcal{M}, \succeq_M)$ is completely ordered, then the ordering $\succeq$ on $\mathcal{X}$ is complete up to equivalence,*** (iv) ***If $(\mathcal{M}, \succeq_M)$ is a lower or upper semi-lattice, then $(\mathcal{X}, \succeq)$ is a lower or upper semi-lattice, respectively, up to equivalence.*** (v) ***A real function f is monotone on $(\mathcal{X}, \succeq)$ if and only if f is a composition of the type***

$$f(x) = \psi(M(x)) = [\psi \circ M](x) \tag{2.1}$$

with ψ a function from the class Ψ of all functions monotone on $(\mathcal{M}, \succeq_M)$.

Proof. For conclusions (i)–(iv) and sufficiency in (v) argue orbit by orbit. The necessity of (v) follows from a result on page 216 of Lehmann (1959), i.e. $f(x) = [\xi \circ M](x)$ for some function ξ on $\mathcal{M}$, together with the monotonicity of f. □

Often there is wide latitude in the choice of an invariant ordering. If $\mathcal{M}$ is a vector space, $\mathcal{K} \subset \mathcal{M}$ a cone, and if $M_1 \succeq_M M_2$ is equivalent to $M_1 - M_2 \in \mathcal{K}$, then $\succeq_M$ is a preordering; $(\mathcal{M}, \succeq_M)$ is partially ordered if and only if $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$ (cf. Wong and Ng (1973)). For cone orderings on $\mathcal{M}$ the functions Ψ monotone on $(\mathcal{M}, \succeq_M)$ are characterized in Marshall, Walkup, and Wets (1967). Elsewhere in this volume Eaton (1984) requires that $x \succeq_E y$ on $\mathcal{X}$ if and only if y lies in the convex hull of the $\mathcal{G}$-orbit of x; given an inner product on $\mathcal{X}$, he characterizes this ordering by quasi-linearization in terms of a maximal invariant function. From the correspondence of orbits to points in $\mathcal{M}$, it is clear that both $\mathcal{X}$ and $\mathcal{M}$ may be ordered using $\succeq_E$, the latter depending only on $\mathcal{G}$ and leaving no latitude for choice. Although that approach and the approach taken here differ, there is common ground as may be seen on comparing our examples with those of Eaton (1984).

We now specialize $\mathcal{X}$ and $\mathcal{G}$ to familiar finite-dimensional linear spaces and the classical groups, respectively. On occasion properties of these spaces support results beyond those of Theorem 1.

3. Orderings on $\mathcal{R}^n$ and $\mathcal{F}_{n \times m}$. Let $\mathcal{R}_+^n$ be the positive orthant of $\mathcal{R}^n$; write $\mathcal{D}_n = \{\mathbf{x} \in \mathcal{R}^n | x_1 \geq x_2 \geq \ldots \geq x_n\}$ and $\mathcal{D}_n^+ = \mathcal{D}_n \cap \mathcal{R}_n^+$; and henceforth consider $\mathcal{F}_{n \times m}$ with $n \geq m$. Denote by $\mathcal{U}(n)$ the unitary $(n \times n)$ matrices and by $\mathcal{S}(n,m)$ the Stiefel manifold in

$\mathcal{F}_{n\times m}$ whose elements satisfy $\mathbf{A}^*\mathbf{A} = \mathbf{I}_m$. The *polar factorization* of $\mathbf{A} \in \mathcal{F}_{n\times m}$ is $\mathbf{A} = \mathbf{LS}$ with $\mathbf{L} \in \mathcal{S}(n,m)$ and $\mathbf{S} = (\mathbf{A}^*\mathbf{A})^{1/2}$ as the Hermitian square root. Its *singular decomposition* is $\mathbf{A} = \mathbf{PD}_\alpha\mathbf{Q}^*$ with $\mathbf{P} \in \mathcal{S}(n,m)$, $\mathbf{Q} \in \mathcal{U}(m)$, and $\mathbf{D}_\alpha = \text{Diag}(\alpha_1, \ldots, \alpha_m)$, a real diagonal matrix of the ordered *singular values* of $\mathbf{A}$, i.e. the non-negative square roots of the characteristic values of $\mathbf{A}^*\mathbf{A}$. Let σ: $\mathcal{F}_{n\times m} \to \mathcal{D}_m^+$ map $\mathbf{A}$ into its ordered singular values.

The following construction is essentially due to von Neumann (1937). For functions γ: $\mathcal{R}^m \to \mathcal{R}^1$ consider the properties

P1. $\gamma(\epsilon_1 x_{i_1}, \ldots, \epsilon_m x_{i_m}) = \gamma(x_1, \ldots, x_m)$, where $\{\epsilon_i = \pm 1; 1 \leq i \leq m)$ and $(i_1, i_2, \ldots, i_m)$ is any permutation of $(1, 2, \ldots, m)$;

P2. If $\{|x_i| \leq |y_i|, 1 \leq i \leq m\}$, then $\gamma(x_1, \ldots, x_m) \leq \gamma(y_1, \ldots, y_m)$ with strict inequality if $|x_i| < |y_i|$ for some i.

P3. $\gamma(x_1, \ldots, x_m) \geq 0$;

P4. $\gamma(cx_1, \ldots, cx_m) = |c|\gamma(x_1, \ldots, x_m)$.

P5. $\gamma(x_1 + y_1, \ldots, x_m + y_m) \leq \gamma(x_1, \ldots, x_m) + \gamma(y_1, \ldots, y_m)$.

Definition 2. Let Γ be the class of functions γ: $\mathcal{R}^m \to \mathcal{R}^1$ having properties P1 and P2, and let $\Gamma_0 \subset \Gamma$ have the additional properties P3, P4 and P5.

Definition 3. Let Φ be the class of functions ϕ: $\mathcal{F}_{n\times m} \to \mathcal{R}^1$ generated by compositions as $\Phi = \{\phi | \phi = \gamma \circ \sigma, \gamma \in \Gamma\}$. Let Φ_0 be the subclass of functions in Φ such that $\phi_0 = \{\phi | \phi = \gamma \circ \sigma, \gamma \in \Gamma_0\}$.

Functions in Γ_0 are the symmetric gauge functions on $\mathcal{R}^m$, and those in Φ_0 are the unitarily invariant norms on $\mathcal{F}_{n\times m}$; von Neumann (1937) showed that these classes generate each other (cf. also Schatten (1970)).

3.1 Orderings on $\mathcal{R}^n$. If $\mathcal{G}$ is the group of 2^n reflections about the coordinate planes in $\mathcal{R}^n$, then the orbit of $\mathbf{x} \in R^n$ has the vertices $\{(\epsilon_1 x_1, \ldots, \epsilon_n x_n); \epsilon_i = \pm 1, 1 \leq i \leq n\}$ of a parallelotope, and a maximal invariant is $M(\mathbf{x}) = (|x_1|, |x_2|, \ldots, |x_n|)$ with range $\mathcal{R}_+^n$. If ordered by coordinates such that $\mathbf{u} \succeq_M \mathbf{v}$ on $\mathcal{R}_+^n$ if and only if $\{u_i \geq v_i; i = 1, \ldots, n\}$, the range $(\mathcal{R}_+^n, \succeq_M)$ is a lattice (cf. Vulikh (1967), for example). The order induced by Definition 1 is that $\mathbf{x} \succeq \mathbf{y}$ on $\mathcal{R}^n$ if and only if $\{|x_i| \geq |y_i|; i = 1, \ldots, n\}$. Theorem 1 now assures that $(\mathcal{R}^n, \succeq)$ is partially ordered symmetrically up to equivalence, and that $(\mathcal{R}^n, \succeq)$ is a lattice up to equivalence under $\mathcal{G}$. For $\mathbf{x},\mathbf{y} \in \mathcal{R}^n$, $\mathbf{x} \wedge \mathbf{y}$ is the orbit identified with $\mathbf{w} \in \mathcal{R}^n$ having $\{w_i = \min(|x_i|, |y_i|); i = 1, \ldots, n\}$, and $\mathbf{x} \vee \mathbf{y}$ is the orbit identified with $\mathbf{z} \in \mathcal{R}^n$ having $\{z_i = \max(|x_i|, |y_i|); i = 1, \ldots, n\}$. By conclusion (v) of Theorem 1 the monotone functions on $(\mathcal{R}^n, \succeq)$ are generated by the class Ψ of functions on $\mathcal{R}_+^n$ increasing (i.e. nondecreasing) in each argument.

Other orderings of interest on $\mathcal{R}^n$ are symmetric under the permutation group $\mathcal{G}$. Then, with $\{x_{(1)} \geq x_{(2)} \geq \ldots \geq x_{(n)}\}$ as the order statistics, $M(\mathbf{x}) = (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ is a maximal invariant on $\mathcal{R}^n$ with range $\mathcal{D}_n$. A fundamental ordering on $\mathcal{R}^n$ is obtained via Theorem 1 through *majorization* on $\mathcal{D}_n$ such that $\mathbf{u} \succeq_M \mathbf{v}$ on $\mathcal{D}_n$ if and only if

$$\Sigma_1^k u_i \geq \Sigma_1^k v_i; k = 1, 2, \ldots . n-1 \tag{3.1}$$

$$\Sigma_1^n u_i = \Sigma_1^n v_i. \tag{3.2}$$

The result is a partial ordering on $\mathcal{R}^n$ up to equivalence. Ordering the elements of $\mathbf{x},\mathbf{y} \in \mathcal{R}^n$ before applying (3.1) and (3.2) is justified formally by (i) of Theorem 1. The functions monotone on $(\mathcal{R}^n, \succeq)$, i.e. the $\mathcal{S}$-convex functions of Schur (1923), may be generated from functions on $\mathcal{D}_n$ and conversely. Part (v) of Theorem 1 thus yields Proposition H.1

of Marshall and Olkin (1979, p. 92) as well as its converse.

Similar remarks apply to a weak majorization on $\mathcal{D}_n$ in which $k = 1, \ldots, n$ in (3.1) and (3.2) is deleted. Examples of functions on $\mathcal{R}^n$ montone under the induced ordering are the symmetric gauge functions (cf. Fan (1951)). Functions monotone under majorization and various weak majorizations are treated in Marshall and Olkin (1979).

3.2 Left-Unitary Ordering on $\mathcal{F}_{n\times m}$. Here $\mathcal{G}$ is the unitary group acting from the left, i.e. $\mathbf{A} \to \mathbf{UA}$ on $\mathcal{F}_{n\times m}$ with $\mathbf{U} \in \mathcal{U}(n)$. A maximal invariant is $M(\mathbf{A}) = \mathbf{A}^*\mathbf{A}$ with range $\mathcal{H}_m^+$ (cf. Vinograde (1950)). The ordering considered here is as follows.

Definition 4. Two matrices $\mathbf{A},\mathbf{B} \in \mathcal{F}_{n\times m}$ are said to be ordered as $\mathbf{A} \succeq \mathbf{B}$ if and only if $\mathbf{A}^*\mathbf{A} \succeq_L \mathbf{B}^*\mathbf{B}$ on $\mathcal{H}_m^+$, i.e. $\mathbf{A}^*\mathbf{A} - \mathbf{B}^*\mathbf{B} \in \mathcal{H}_m^+$. The ordering $\mathbf{A} \succ \mathbf{B}$ is strict whenever $\mathbf{A}^*\mathbf{A} - \mathbf{B}^*\mathbf{B}$ is positive definite.

Basic properties of $(\mathcal{F}_{n\times m}, \succeq)$, ordered as in Definition 4, are given next, where $(\mathcal{H}_m^+, \succeq_L)$ is used in lieu of $(\mathcal{M}, \succeq_M)$.

THEOREM 2. *Let $\succeq$ be a relation on $\mathcal{F}_{n\times m}$ induced from $(\mathcal{H}_m^+, \succeq_L)$ as in Definition 4.* (i) *The relation $\succeq$ is invariant in the sense that $\mathbf{A} \succeq \mathbf{B}$ on $\mathcal{F}_{n\times m}$ if and only if $\mathbf{PAC} \succeq \mathbf{QBC}$ for any $\mathbf{P},\mathbf{Q} \in \mathcal{U}(n)$ and $\mathbf{C} \in Gl(m)$.* (ii) *$(\mathcal{F}_{n\times m}, \succeq)$ is partially ordered up to equivalence under $\mathcal{G}$.* (iii) *If $m = 1$ the ordering $\succeq$ is complete up to equivalence.* (iv) *If $\mathbf{A} \succeq \mathbf{B}$ on $\mathcal{F}_{n\times m}$, then $\mathbf{AG} \succeq \mathbf{BG}$ on $\mathcal{F}_{n\times s}$ for any $\mathbf{G} \in \mathcal{F}_{m\times s}$.* (v) *For $\mathbf{A},\mathbf{B} \in \mathcal{F}_{n\times m}$ a necessary and sufficient condition that $\mathbf{A} + \mathbf{B} \succeq \mathbf{A}$ is that $\mathbf{B}^*\mathbf{B} + \mathbf{A}^*\mathbf{B} + \mathbf{B}^*\mathbf{A} \in \mathcal{H}_m^+$. In particular, if $\mathbf{A}^*\mathbf{B} = \mathbf{0}$, then $\mathbf{A} + \mathbf{B} \succeq \mathbf{A}$.*

Proof. Conclusions (i)–(iii) follow from their antecedents in Theorem 1 together with properties of $(\mathcal{H}_m^+, \succeq_L)$; conclusion (iv) is immediate; and (v) follows on expanding $(\mathbf{A} + \mathbf{B})^*(\mathbf{A} + \mathbf{B})$. □

Claims for lattice properties of $(\mathcal{F}_{n\times m}, \succeq)$ are not available through Theorem 1. Halmos (1958, p. 142) showed that $(\mathcal{H}_m, \succeq_L)$ is not a lattice.

Well known properties of $(\mathcal{H}_n, \succeq_L)$ are 1) if $\mathbf{A} \succeq_L \mathbf{B} \succ_L \mathbf{0}$, then $\mathbf{B}^{-1} \succeq_L \mathbf{A}^{-1} \succ_L \mathbf{0}$ (cf. Loewner (1934)) and 2) if $\mathbf{A} \succeq_L \mathbf{B}$, then their ordered characteristic values, $Ch(\mathbf{A}) = \{\alpha_1 \geq \ldots \geq \alpha_n\}$ and $Ch(\mathbf{B}) = \{\beta_1 \geq \ldots \geq \beta_n\}$, satisfy $\{\alpha_i \geq \beta_i;\ 1 \leq i \leq n\}$ (cf. Bellman (1960), p. 115). The first was extended to singular matrices in $\mathcal{H}_n^+$ by Milliken and Akdeniz (1977) using the pseudo-inverse of Moore (1920) and Penrose (1955). Thus these inverse operators are order-reversing on the boundary and interior of $\mathcal{H}_n^+$ under $\succeq_L$. Corresponding properties are shown next for $(\mathcal{F}_{n\times m}, \succeq)$ in terms of 1′) the Moore-Penrose inverse operator $\mathbf{A} \to \mathbf{A}^\dagger$ on $\mathcal{F}_{n\times m}$ and 2′) the singular-value mapping $\mathbf{A} \to \sigma(\mathbf{A})$.

THEOREM 3. *Let $\sigma(\mathbf{A}) = \{\alpha_1 \geq \ldots \geq \alpha_m\}$ and $\sigma(\mathbf{B}) = \{\beta_1 \geq \ldots \geq \beta_m\}$ be the ordered singular values of $\mathbf{A},\mathbf{B} \in \mathcal{F}_{n\times m}$, and let $\mathbf{A}^\dagger$ and $\mathbf{B}^\dagger$ be their respective Moore-Penrose inverses.* (i) *If $\mathbf{A} \succeq \mathbf{B}$, then $\alpha_i \geq \beta_i$ for $i = 1, \ldots, m$.* (ii) *If $\mathbf{A} \succeq \mathbf{B}$ and if $\mathbf{A}$ and $\mathbf{B}$ have rank $s \leq m$, then $(\mathbf{B}^\dagger)^* \succeq (\mathbf{A}^\dagger)^*$ on $\mathcal{F}_{n\times m}$.*

Proof. Conclusion (i) is a restatement of property 2) of $(\mathrm{H}_m^+, \succeq_L)$. Conclusion (ii) follows because $\mathbf{A} \succeq \mathbf{B}$ implies $\mathbf{A}^*\mathbf{A} \succeq_L \mathbf{B}^*\mathbf{B}$, which implies (cf. Milliken and Akdeniz (1977)) that $(\mathbf{B}^*\mathbf{B})^\dagger \succeq_L (\mathbf{A}^*\mathbf{A})^\dagger$, which in turn implies $\mathbf{B}^\dagger(\mathbf{B}^\dagger)^* \succeq_L \mathbf{A}^\dagger(\mathbf{A}^\dagger)^*$ and thus (ii). The last step uses the singular decomposition $\mathbf{F} = \mathbf{PDQ}^*$ with $\mathbf{P} \in \mathcal{U}(n)$, $\mathbf{Q} \in \mathcal{U}(m)$, and $\mathbf{D}(n\times m)$, and the fact that $\mathbf{F}^\dagger = \mathbf{QD}^\dagger\mathbf{P}^*$. □

The functions monotone on $(\mathcal{F}_{n\times m}, \succeq)$ may be generated by composition as $\{f(\mathbf{X}) = \psi(\mathbf{X}^*\mathbf{X});\ \psi \in \Psi\}$, with Ψ the real functions monotone on $(\mathcal{H}_m^+, \succeq_L)$ as characterized by Marshall, Walkup, and Wets (1967). Other results follow.

THEOREM 4. *Let Φ be the class of functions on $(\mathcal{F}_{n\times m}, \succeq)$ as in Definition 3. If $\phi \in \Phi$, then ϕ is monotone. In particular, any $\phi \in \Phi_0$ is a monotone norm.*

Proof. That ϕ is order-preserving follows on combining conclusion (i) of Theorem 3 with property (ii) of Definition 2, for $\mathbf{A} \succeq \mathbf{B}$ implies $\phi(\mathbf{A}) = \gamma(\sigma_1(\mathbf{A}), \ldots, \sigma_m(\mathbf{A})) \geq \gamma(\sigma_1(\mathbf{B}), \ldots, \sigma_m(\mathbf{B})) = \phi(\mathbf{B})$. In particular, any $\phi \in \Phi_0$, a norm, is a monotone norm. □

In conclusion, note that the ordering $\succeq$ of Definition 4 is an alternative to Loewner's ordering $\succeq_L$ on H_n; clearly the two orderings coincide on $\mathcal{H}_n^+$. Examples of elements in $\mathcal{H}_2$ ordered by $\succeq$ but not $\succeq_L$ are the diagonal matrices $\mathbf{A}$ = Diag(2,–2) and $\mathbf{B}$ = Diag(1, -1).

3.3 Unitary Orderings on $\mathcal{F}_{n\times m}$. Let $\mathcal{G}$ be the unitary group on $\mathcal{F}_{n\times m}$ taking $\mathbf{A}$ into $\mathbf{UAV}$ with $\mathbf{U} \in \mathcal{U}(n)$ and $\mathbf{V} \in \mathcal{U}(m)$. A maximal invariant is $\sigma(\mathbf{A})$ with range $\mathcal{D}_m^+$, each orbit in $\mathcal{F}_{n\times m}$ having matrices with the same ordered singular values.

Orderings on $\mathcal{D}_m^+$ of interest here are those of Section 3.1: (i) ordering by coordinates, (ii) majorization, and (iii) weak majorization. Properties of $(\mathcal{F}_{n\times m}, \succeq)$ under the induced orderings follow from Theorem 1 as before. Functions monotone on $(\mathcal{F}_{n\times m}, \succeq)$ are equivalent to compositions $\{f(\mathbf{A}) = \psi(\sigma(\mathbf{A}));\ \psi \in \Psi\}$ where, for the three cases, (i) Ψ consists of functions on $\mathcal{D}_m^+$ increasing in each argument, (ii) Ψ consists of $\mathcal{S}$-convex functions, and (iii) Ψ is the class Γ_0 of Definition 2. The latter combines Theorem 1(v) with a result of Fan (1951). For this case the class of monotone functions is Φ_0, the unitarily invariant norms being monotone on $(\mathcal{F}_{n\times m}, \succeq)$ when ordering is induced through weak majorization of the singular values. For case (i) $(\mathcal{F}_{n\times m}, \succeq)$ is a lattice up to equivalence under $\mathcal{G}$.

PART II. APPLICATIONS

1. Introduction. The foregoing concepts apply in a variety of settings, where different orderings serve different purposes and a careful choice may yield results not otherwise attainable. Unless stated otherwise, we take $(\mathcal{F}_{n\times m}, \succeq)$ to be ordered as in Definition 4.

Subsequently define $\mathcal{F}_{n\times m}$ over $\mathcal{R}^1$ and let $\mathcal{O}(n)$ be the group of real orthogonal $(n\times n)$ matrices. Let $S_{n,m}(\Theta, \Gamma\times\Xi)$ be the class of ellipsoidal matrix distributions on $\mathcal{F}_{n\times m}$ with typical density

$$f(\mathbf{Y}) = g(tr(\mathbf{Y}-\Theta)\Xi^{-1}(\mathbf{Y}-\Theta)'\Gamma^{-1}); \tag{1.1}$$

let $U_{n,m}(\Theta, \Gamma\times\Xi)$ be the subclass of distributions unimodal in the sense of Anderson (1955); let $G_{n,m}(\Theta, \Gamma\times\Xi)$ be the subclass of these consisting of scale mixtures of matrix Gaussian laws; and let $L_{n,m}(\Theta, \Lambda)$ be the distribution on $\mathcal{F}_{n\times m}$ with typical density

$$f(\mathbf{Y}) = h(\Lambda'(\mathbf{Y}-\Theta)'(\mathbf{Y}-\Theta)\Lambda). \tag{1.2}$$

Here $\Theta \in \mathcal{F}_{n\times m}$ consists of location parameters, while $\Xi \in \mathcal{H}_m^+$, $\Gamma \in \mathcal{H}_n^+$, and $\Lambda \in \mathcal{F}_{m\times m}$ are scale parameters, all nonsingular. These distributions are considered in Jensen and Good (1981). When $m = 1$ these specialize to the classes $S_n(\theta, \Gamma)$, $U_n(\theta, \Gamma)$, and $G_n(\theta, \Gamma)$ on $\mathcal{R}^n$. The distribution of $\mathbf{W}$ is denoted by $\mathcal{L}(\mathbf{W})$.

A useful ordering for probability measures is the following (compare Sherman (1955)), where $\mathcal{X}$ is a linear space having the zero element $0 \in \mathcal{X}$.

Definition 5. Let $(\mathcal{X}, \mathcal{B}, \cdot)$ be a measurable space. The probability measure μ is *more peaked* about $0 \in \mathcal{X}$ than ν if $\mu(A) \geq \nu(A)$ for every set in the class τ of compact convex measurable sets A symmetric under reflection, i.e. $x \in A$ implies $-x \in A$. Denote this ordering by $\mu \succeq_p \nu$.

2. Peakedness of Measures. For suitable measures μ and ν on $\mathcal{R}^n$ symmetric about $\mathbf{0}$ and having the dispersion matrices Σ_μ and Σ_ν such that $\Sigma_\nu \succeq_L \Sigma_\mu$, μ is more peaked about $\mathbf{0}$ than ν. This was shown for Gaussian measures on $\mathcal{R}^n$ by Anderson (1955) and for $S_n(\mathbf{0},\Gamma)$ by Fefferman, Jodeit, and Perlman (1972). We extend these results to linear transformations from $\mathcal{R}^n$ to $\mathcal{R}^m$ with $m \leq n$, and we also supply a converse.

THEOREM 5. *Let μ_A and μ_B be probability measures on $\mathcal{R}^m$ induced by $\mathbf{y} \to \mathbf{A}'\mathbf{y}$ and $\mathbf{y} \to \mathbf{B}'\mathbf{y}$ from $\mathcal{L}(\mathbf{y}) \in S_n(\mathbf{0},\mathbf{I}_n)$ with $\mathbf{A},\mathbf{B} \in \mathcal{F}_{n\times m}$. Then $\mu_A \succeq_p \mu_B$ on $\mathcal{R}^m$ if and only if $\mathbf{B} \succeq \mathbf{A}$ on $(\mathcal{F}_{n\times m}, \succeq)$.*

Proof. (i) Clearly $\mathbf{A}'\mathbf{A} = \Sigma_A$ and $\mathbf{B}'\mathbf{B} = \Sigma_B$ are the scale parameters of μ_A and μ_B. If $\mathbf{B} \succeq \mathbf{A}$, then $\Sigma_B \succeq_L \Sigma_A$ and the ordering $\mu_A \succeq_p \mu_B$ follows from that of Fefferman, Jodeit, and Perlman (1972). (ii) Conversely, suppose that $\mu_A \succeq_p \mu_B$ but neither $\mathbf{A} \succeq \mathbf{B}$ nor $\mathbf{B} \succeq \mathbf{A}$. As $\succeq_p$ is preserved under nonsingular linear transformations, simultaneously reduce Σ_A to $\mathbf{I}_m$ and Σ_B to $\mathbf{D} = \text{Diag}(\delta_1, \ldots, \delta_m)$ where, for some $k \in (1,m)$, $\delta_1 \leq \ldots \leq \delta_k < 1 \leq \delta_{k+1} \leq \ldots \leq \delta_m\}$. Then the marginal measures ν_A and ν_B on $\mathcal{R}^k$ are ordered as $\nu_B \succeq_p \nu_A$, and necessity follows by contradiction using the fact that probability measures on $\mathcal{R}^n$ are tight. □

3. Linear Estimation. The problem is to estimate Θ in the matrix model $\mathbf{Y} = \mathbf{X}\Theta + \mathbf{E}$ with $\mathbf{Y} \in \mathcal{F}_{n\times m}$ observable, $\mathbf{X} \in \mathcal{F}_{n\times r}$ known of rank $r \leq n$, and $\mathbf{E} \in \mathcal{F}_{n\times m}$ a matrix of random errors. Minimizing $Q(\Theta) = tr(\mathbf{Y}-\mathbf{X}\Theta)'(\mathbf{Y}-\mathbf{X}\Theta)$ as Θ varies yields the least-squares solution $\hat{\Theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ for Θ and $\mathbf{X}\hat{\Theta}$ for approximating $\mathbf{Y}$. We show much stronger minimizing properties.

THEOREM 6. *Suppose $\mathbf{Y} = \mathbf{X}\Theta + \mathbf{E}$ and order $(\mathcal{F}_{n\times m}, \succeq)$ as in Definition 4.* (i) $\hat{\Theta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ *is minimizing on $(\mathcal{F}_{n\times m}, \succeq)$ in the sense that $(\mathbf{Y}-\mathbf{X}\Theta) \succeq (\mathbf{Y}-\mathbf{X}\hat{\Theta})$ for every $\Theta \in \mathcal{F}_{r\times m}$.* (ii) *$\psi(\mathbf{Y}-\mathbf{X}\hat{\Theta}) \leq \psi(\mathbf{Y}-\mathbf{X}\Theta)$ for every ψ in the class Ψ of functions monotone on $(\mathcal{F}_{n\times m}, \succeq)$.* (iii) *$\hat{\Theta}$ is the minimum-norm solution to $\min_{\Theta\in\mathcal{F}_{r\times m}} \|\mathbf{Y}-\mathbf{X}\Theta\|_\Phi$ for every unitarily invariant norm $\|\cdot\|_\Phi$ on $\mathcal{F}_{n\times m}$.*

Proof. Because $\mathbf{X}'\mathbf{X}\hat{\Theta} - \mathbf{X}'\mathbf{Y} = \mathbf{0}$, the expansion

$$(\mathbf{Y}-\mathbf{X}\Theta)'(\mathbf{Y}-\mathbf{X}\Theta) = (\mathbf{Y}-\mathbf{X}\hat{\Theta})'(\mathbf{Y}-\mathbf{X}\hat{\Theta}) + (\hat{\Theta}-\Theta)'\mathbf{X}'\mathbf{X}(\hat{\Theta}-\Theta)$$

yields conclusion (i) directly. Conclusion (ii) follows by monotonicity, and conclusion (iii) from Theorem 4, where the unitarily invariant norms are shown to be monotone. □

Conclusion (iii) was obtained by Rao (1980) as a consequence of ordering the singular values $\sigma(\mathbf{Y}-\mathbf{X}\Theta)$ and $\sigma(\mathbf{Y}-\mathbf{X}\hat{\Theta})$. Our conclusion (i) is stronger, as it implies Rao's ordering using Theorem 3.

4. Ordered Designs. Specialize the foregoing model with $m = 1$ to $\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$ and consider the choice of design as it pertains to testing $\mathbf{H}$: $\beta = \mathbf{0}$ against A: $\beta \neq \mathbf{0}$. If $\mathcal{L}(\mathbf{e}) \in U_n(\mathbf{0},\sigma^2\mathbf{I}_n)$, then the power function of the normal-theory test for H against A depends on the parameters (β,σ^2) and the design $\mathbf{X}$ only through $\lambda = \beta'\mathbf{X}'\mathbf{X}\beta/\sigma^2$, and it increases monotonically with λ; cf. Theorem 3.2 of Jensen (1979). Suppose one of the designs $\mathbf{A},\mathbf{B} \in \mathcal{F}_{n\times r}$ is to be chosen. A connection between the ordering $(\mathcal{F}_{n\times r}, \succeq)$ of Definition 4 and the power of this test is given in the following theorem for any unimodal spherical law of errors.

THEOREM 7. *Suppose* $\mathcal{L}(\mathbf{y}) \in U_n(\mathbf{X}\beta, \sigma^2\mathbf{I}_n)$ *with* $\mathbf{X} \in \mathcal{F}_{n\times r}$ *a design matrix of rank* $r \le n$. *Of two designs* $\mathbf{A},\mathbf{B} \in \mathcal{F}_{n\times r}$, *the F-test at level* α *for testing* $\mathbf{H}$: $\beta = \mathbf{0}$ *against* A:$\beta \neq \mathbf{0}$ *is uniformly more powerful using Design A than using Design B if and only if* $\mathbf{A} \succeq \mathbf{B}$ *on* $\mathcal{F}_{n\times r}$.

Proof. Sufficiency follows from the preceding paragraph because the power function depends only on λ. A proof for necessity parallels that of Theorem 5 on identifying subspaces where each is more powerful when $\mathbf{A} \not\succeq \mathbf{B}$. □

5. The T² Chart under Scale Ordering. Hotelling's (1947) T^2 chart uses samples from a vector-valued process to monitor the stationarity of means over time. Its *run length* is the number of successive samples taken before signaling that the process is not in control. Suppose successive observations are independent Gaussian vectors on $\mathcal{R}^m$ having parameters (μ,Σ), with Σ characteristic of the process. In practice efforts are made to tighten the process to reduce its variability. The effects of such reduction on run lengths are as follows.

THEOREM 8. *Let* N_1 *be the run length of Hotelling's (1947)* T^2 *chart for monitoring a stationary Gaussian process with parameters* (μ, Σ_1), *and let* N_2 *be the run length of* T^2 *for a tightened process with parameters* (μ, Σ_2). *Then* N_2 *is stochastically smaller than* N_1 *for all* $\mu \in \mathcal{R}^m$ *if and only if* $\Sigma_1 \succeq_L \Sigma_2$ *on* $(\mathcal{H}_m^+, \succeq_L)$.

Proof. The proof parallels that of Theorem 7. □

6. Canonical Analysis. Let $\mathbf{z} = (\mathbf{z}_1', \mathbf{z}_2')'$ be Gaussian on $\mathcal{R}^n$ having zero means and the dispersion matrix

$$\Omega = \begin{pmatrix} \mathbf{I}_r & \mathbf{R} \\ \mathbf{R}' & \mathbf{I}_s \end{pmatrix}, \mathbf{R} \in \mathcal{F}_{r\times s}, r \le s, \tag{6.1}$$

and consider the quadratic forms $U_1 = \mathbf{z}_1'\mathbf{z}_1$ and $U_2 = \mathbf{z}_2'\mathbf{z}_2$. Their joint distribution depends only on $\sigma(\mathbf{R})$, the canonical correlations of Hotelling (1936). If an ordering on $(\mathcal{F}_{r\times s}, \succeq)$ is induced by coordinate-wise ordering of the singular values on $\mathcal{D}_r^+$, then a monotonicity property of certain probabilities is given in the following.

THEOREM 9. *Let* $\mu_R(\cdot)$ *be the joint measure of* $U_1 = \mathbf{z}_1'\mathbf{z}_1$ *and* $U_2 = \mathbf{z}_2'\mathbf{z}_2$ *on* $\mathcal{R}_+^2$ *having the cross correlation matrix* $\mathbf{R}$ *in (6.1) with* $r = s$. *Then for every measurable set* $A \subset \mathcal{R}_+^1$, *the measure* $\mu_R(A\times A)$ *is order-preserving in the sense that* $\Xi \succeq \Gamma$ *on* $(\mathcal{F}_{r\times r}, \succeq)$ *implies that* $\mu_\Xi(A\times A) \geq \mu_\Gamma(A\times A)$.

Proof. An expansion in the Lancaster (1958) canonical form was given by Jensen (1970) for the joint distribution, the coefficients $G_k(\rho)$ depending on $\rho = \sigma(\mathbf{R})$. The proof consists of integrating the expansion over $A\times A$ and showing that the resulting expression is an increasing function of ρ. □

We next show that a bivariate Chebyshev bound is monotone when considered as a function of $\mathbf{R}$ on $(\mathcal{F}_{r\times s}, \succeq)$ for any $r \le s$. This result is essentially distribution-free. Define

$$B(\delta_1, \delta_2;\mathbf{R}) = ((s-r)/\delta_2) + \Sigma_{i=1}^r\{(\delta_1+\delta_2) + [(\delta_1+\delta_2)^2 - 4\rho_i^2\delta_1\delta_2]^{1/2}\}/2\delta_1\delta_2 \tag{6.2}$$

in terms of the canonical correlations $\{\rho_1, \ldots, \rho_r\}$ of $\mathbf{y}_1$ and $\mathbf{y}_2$ for fixed δ_1 and δ_2.

THEOREM 10. *Let* $\mathbf{y} = (\mathbf{y}'_1, \mathbf{y}'_2)'$ *be a random vector of order* $r + s = n$, *with* $r \leqslant s$, *having the mean* $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1, \boldsymbol{\mu}'_2)'$, *the dispersion matrix* $\Sigma = [\Sigma_{ij}]$, *and* $\mathbf{R} = \Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2}$. *Then* (i) $P((\mathbf{y}_1-\boldsymbol{\mu}_1)'\Sigma_{11}^{-1}(\mathbf{y}_1-\boldsymbol{\mu}_1) \leqslant \delta_1, (\mathbf{y}_2-\boldsymbol{\mu}_2)'\Sigma_{22}^{-1}(\mathbf{y}_2-\boldsymbol{\mu}_2) \leqslant \delta_2) \geqslant 1-B(\delta_1,\delta_2;\mathbf{R})$; (ii) $B(\delta_1,\delta_2;\mathbf{R})$ *is monotone decreasing on* $(\mathcal{F}_{r\times s}, \succeq)$.

Proof. Conclusion (i) is given in Jensen (1982). Conclusion (ii) follows from expression (6.2), from the mapping $\sigma(\mathbf{R}) = (\rho_1, \ldots, \rho_r)$, and from the fact that $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1/2} \succeq \Omega_{11}^{-1/2}\Omega_{12}\Omega_{22}^{-1/2}$ on $(\mathcal{F}_{r\times s}, \succeq)$, for example, if and only if their singular values are pairwise ordered on $\mathcal{D}_r^+$. □

Observe that Theorems 9 and 10 remain valid when $(\mathcal{F}_{r\times s}, \succeq)$ is ordered as in Definition 4. This follows from conclusion (i) of Theorem 3.

7. Signal Detection under Symmetry. Each channel of a k-channel receiver accepts an input vector $\mathbf{y} \in \mathcal{R}^n$ that is either processed as a signal or suppressed as noise depending on whether the input amplitude $||\mathbf{y}|| = (\mathbf{y}'\mathbf{y})^{1/2}$ does or does not exceed a threshold value c. Thus for k channels $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_k]$ represents the input, typically correlated between channels, and a probabilistic assessment of the system performance focuses on expressions of the type

$$F(c_1, \ldots, c_k; \Theta) = P_\Theta(||\mathbf{y}_1|| \leq c_1, \ldots, ||\mathbf{y}_k|| \leq c_k), \tag{7.1}$$

when signals of varying strengths actually enter the system. A useful ordering is the following.

THEOREM 11. *Suppose* $\mathcal{L}(\mathbf{Y}) \in L_{n,k}(\Theta,\Lambda)$ *and its pdf is unimodal. Then the probability*

$$F(c_1, \ldots, c_k;\Theta) = P_\Theta(||\mathbf{y}_1|| \leq c_1, \ldots, ||\mathbf{y}_k|| \leq c_k),$$

when considered as a function of Θ *with* $\{c_1, \ldots, c_k\}$ *fixed, is monotone decreasing on* $(\mathcal{F}_{n\times k}, \succeq)$ *under the ordering of Definition 4.*

Proof. Let $A \subset \mathcal{F}_{n\times k}$ be the set

$$A = \{\mathbf{Y} \in \mathcal{F}_{n\times k} | \mathbf{y}'_1\mathbf{y}_1 \leq c_1^2, \ldots, \mathbf{y}'_k\mathbf{y}_k \leq c_k^2\};$$

Let $\mathcal{G}$ be the group $\mathcal{O}(n)$ acting on $\mathbf{Y}$ from the left; and observe that (i) A is a convex $\mathcal{G}$-invariant subset of $\mathcal{F}_{n\times k}$ and (ii) the pdf of $(\mathbf{Y}-\Theta)$ is a nonnegative real-valued, $\mathcal{G}$-invariant and unimodal function on $\mathcal{F}_{n\times k}$. Conditions (i) and (ii) satisfy the requirements of Theorem 5 of Mudholkar (1966) which assures that

$$P_\Theta(A) \geq P_{\Theta_0}(A) \tag{7.2}$$

for any Θ in the convex hull of the $\mathcal{G}$-orbit of Θ_0. This orbit is characterized by constant values of the maximal invariant function $\Theta'_0\Theta_0$. On taking sections, we infer that Θ is in the convex hull of the $\mathcal{G}$-orbit of Θ_0 if for all $\mathbf{a} \in \mathcal{R}^k$, $a'\Theta'\Theta a \leq a'\Theta'_0\Theta_0 a$, i.e. if $\Theta'_0\Theta_0 \succeq_L \Theta'\Theta$. But from Definition 4, this fact together with (7.2) are equivalent to the assertion of the theorem. □

In practice this assures that the larger the shift in the sense of the ordering $\succeq$, the greater the probability that signals are correctly identified and processed as signals in one or more channels. In particular, in back-up systems designed with redundancies, the detection probability will increase with the magnitude of the signal.

Acknowledgment. Thanks are due a referee for suggestions toward revision.

REFERENCES

ANDERSON, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc. 6* 170–176.

BELLMAN, R. (1960). *Introduction to Matrix Analysis.* McGraw-Hill, New York.

EATON, M. L. (1984). On group induced orderings, monotone functions, and convolution theorems. In *Inequalities in Statistics and Probability*, Y. L. Tong, ed., Institute of Mathematical Statistics, Hayward, Ca., 13–25.

FAN, K. (1951) Maximum properties and inequalities for the eigenvalues of completely continuous operators. *Proc. Nat. Acad. Sci. U. S. A. 37* 760–766.

FEFFERMAN, C., JODEIT, M., and PERLMAN, M. D. (1972). A spherical surface measure inequality for convex sets. *Proc. Amer. Math. Soc. 33* 114–119.

HALMOS, P. (1958) *Finite-Dimensional Vector Spaces,* Second. ed., Van Nostrand, New York.

HARTWIG, R. E. (1976). Block generalized inverses. *Arch. Rat. Mech. Anal. 61* 197–251.

HOTELLING, H. (1947). Multivariate quality control. In *Techniques of Statistical Analysis*, C. Eisenhart, M. W. Hastay, and W. A. Wallis, eds., McGraw-Hill, New York, 111–184.

JENSEN, D. R. (1970). The joint distribution of quadratic forms and related distributions. *Australian J. Statist. 12* 13–22.

JENSEN, D. R. (1979). Linear models without moments. *Biometrika 66* 611–617.

JENSEN, D. R. (1982). Inequalities for joint distributions of quadratic forms. *SIAM J. Appl. Math. 42* 297–301.

JENSEN, D. R. and GOOD, I. J. (1981). Invariant distributions associated with matrix laws under structural symmetry. *J. Roy. Statist. Soc. B 43* 327–332.

LANCASTER, H. O. (1958). The structure of bivariate distributions. *Ann. Math. Statist. 29* 719–736.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

LOEWNER, C. (1934). Über monotone Matrixfunktionen. *Math. Z. 38* 177–216.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic Press, New York.

MARSHALL, A. W., WALKUP, D. W., and WETS, R. J. B. (1967). Order-preserving functions; applications to majorization and order statistics. *Pacific J. Math. 23* 569–584.

MILLIKEN, G. A. and AKDENIZ, F. (1977). A theorem on the difference of generalized inverses of two nonnegative matrices. *Comm. Statist. A6* 73–79.

MOORE, E. H. (1920). On the reciprocal of the general algebraic matrix (abstract). *Bull. Amer. Math. Soc. 26* 394–395.

MUDHOLKAR, G. S. (1966). The integral of an invariant unimodal function over an invariant convex set—An inequality and applications. *Proc. Amer. Math. Soc. 17* 1327–1333.

NEUMANN, J. von (1937). Some matrix inequalities and metrization of matric space. *Tomsk Univ. Rev. 1* 286–300.

PENROSE, R. (1955). A generalized inverse for matrices. *Proc. Cambridge Philos. Soc. 51* 406–413.

RAO, C. R. (1980). Matrix approximations and reduction of dimensionality in multivariate statistical analysis. In *Multivariate Analysis—V*, P. R. Krishnaiah, ed., North Holland, 3–22.

SCHATTEN, R. (1970). *Norm Ideals of Completely Continuous Operators.* Springer-Verlag, Berlin and New York.

SCHUR, I. (1923). Über eine Klasse von Mittelbildungen mit Anwendungen die Determinantentheorie Sitzungsber. *Berlin Math. Gesellschaft 22* 9–20.

SHERMAN, S. (1955). A theorem on convex sets with applications. *Ann. Math. Statist. 26* 763–766.

VINOGRADE, B. (1950). Canonical positive definite matrices under internal linear transformations. *Proc. Amer. Math. Soc. 1* 159–161.

VULIKH, B. Z. (1967). *Introduction to the Theory of Partially Ordered Spaces.* Wolters-Noordhoff, Groningen.

WONG, Y.-C. and NG, K.-F. (1973). *Partially Ordered Topological Vector Spaces.* Clarendon, Oxford.