

AD-A097 110

STANFORD UNIV CA DEPT OF STATISTICS

F/6 12/1

A NEW PROOF OF ADMISSIBILITY OF TESTS IN THE MULTIVARIATE ANALY--ETC(U)

JAN 81 T W ANDERSON; A TAKEMURA

N00014-75-C-0442

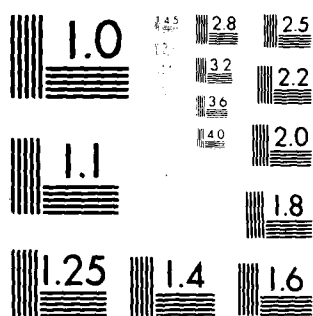
UNCLASSIFIED

TR-48

NL

[X]  
A-111-0

END  
DATE  
FILMED  
3-84  
DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

12

AD AUG 11 1981

(6) A NEW PROOF OF ADMISSIBILITY OF TESTS IN THE  
MULTIVARIATE ANALYSIS OF VARIANCE.

BY

(1) T. W. ANDERSON and AKIMICHI TAKEMURA

(1) TECHNICAL REPORT NO. 48  
JANUARY 1981

DTIC  
ELECTE  
MAR 3 1 1981  
S D E

(5)  
PREPARED UNDER CONTRACT N00014-75-C-0442  
(NR-042-034)

OFFICE OF NAVAL RESEARCH

THEODORE W. ANDERSON, PROJECT DIRECTOR

(14) MK 42

12 /

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

DTIC FILE COPY



DISTRIBUTION STATEMENT A

33252

81 3 30 077

A NEW PROOF OF ADMISSIBILITY OF TESTS IN THE  
MULTIVARIATE ANALYSIS OF VARIANCE

by

T. W. ANDERSON

and

AKIMICHI TAKEMURA

TECHNICAL REPORT NO. 48

JANUARY 1981

PREPARED UNDER CONTRACT N00014-75-C-0442  
(NR-042-034)  
OFFICE OF NAVAL RESEARCH

Theodore W. Anderson, Project Director

Reproduction in Whole or in Part is Permitted for  
any Purpose of the United States Government.  
Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA

|                    |                                            |
|--------------------|--------------------------------------------|
| Accession For      |                                            |
| NTIS CBA&I         | <input checked="checked" type="checkbox"/> |
| DTIC TAB           | <input type="checkbox"/>                   |
| Unannounced        | <input type="checkbox"/>                   |
| Justification      |                                            |
|                    |                                            |
| Dist               |                                            |
| Distribution Codes |                                            |
|                    |                                            |
| Dist               | Special                                    |
| A                  |                                            |

A NEW PROOF OF ADMISSIBILITY  
OF TESTS IN THE MULTIVARIATE ANALYSIS OF VARIANCE

T. W. Anderson and Akimichi Takemura

1. Introduction.

A general theorem on the admissibility of tests of the general multivariate linear hypothesis was proved by Schwartz (1967) and Ghosh (1964) using Stein's theorem (1956). In using Stein's theorem there are two conditions to prove: (i) convexity of the acceptance region and (ii) existence of certain alternative hypotheses. In his paper on the admissibility of Hotelling's  $T^2$ -test Stein (1956) proved the convexity of the acceptance region by showing that it is an intersection of half-spaces. Schwartz (1967) and Ghosh (1964) followed this approach.

The purpose of this paper is to present a new proof of the admissibility of the tests of the general multivariate linear hypothesis. We show the convexity of the acceptance region more directly; that is, a convex combination of two sample points in the acceptance region again belongs to the region. The separation of conditions (i) and (ii) simplifies considerably the proof of the convexity condition (i) and makes its geometrical meaning clearer. We shall be explicit also in proving the condition (ii).

In Section 2 we state the admissibility results in several forms and discuss their relations. In Section 3 Stein's theorem on the admissibility of tests in the general exponential family framework is stated. The rest of the paper is devoted to the proof of the theorems in Section 2.

## 2. The Problem and Main Results.

In this section we set up the problem and state the admissibility results in several forms. Discussion of the relations of those forms will be given. For proofs see Section 4.

The problem of testing the general multivariate linear hypothesis can be written in the following canonical form. Let  $\underline{X}$  ( $p \times m$ ),  $\underline{Y}$  ( $p \times r$ ), and  $\underline{Z}$  ( $p \times n$ ) be random matrices such that their columns are independently normally distributed with a common covariance matrix  $\underline{\Sigma}$  and means

$$(2.1) \quad \mathcal{E}\underline{X} = \underline{M}, \quad \mathcal{E}\underline{Y} = \underline{N}, \quad \mathcal{E}\underline{Z} = \underline{0},$$

respectively. (See Anderson (1958), sec. 8.11.) The null hypothesis is  $H_0: \underline{M} = \underline{0}$  and the alternative hypothesis is  $\underline{M} \neq \underline{0}$ . The parameter space  $\Omega$  is given by

$$(2.2) \quad \Omega = \{(\underline{M}, \underline{N}, \underline{\Sigma}) \mid \underline{\Sigma} \text{ positive definite} \}.$$

The null hypothesis is

$$(2.3) \quad \Omega_0 = \{(\underline{M}, \underline{N}, \underline{\Sigma}) \mid \underline{M} = \underline{0}, \underline{\Sigma} \text{ positive definite} \}.$$

A test  $T^*$  of the null hypothesis  $H_0: \underline{\omega} \in \Omega_0$  is said to be admissible if there exists no other test  $T$  such that

$$(2.4) \quad \begin{aligned} \Pr\{\text{Reject } H_0 \mid T, \underline{\omega}\} &\leq \Pr\{\text{Reject } H_0 \mid T^*, \underline{\omega}\}, \quad \underline{\omega} \in \Omega_0, \\ \Pr\{\text{Reject } H_0 \mid T, \underline{\omega}\} &\geq \Pr\{\text{Reject } H_0 \mid T^*, \underline{\omega}\}, \quad \underline{\omega} \in \Omega - \Omega_0, \end{aligned}$$

with strict inequality for at least one  $\underline{\omega}$ .

The usual tests of the above null hypothesis can be given in terms of the (nonzero) roots of the following determinantal equation:

$$(2.5) \quad |\underline{\underline{XX'}} - \lambda(\underline{\underline{ZZ'}} + \underline{\underline{XX'}})| = |\underline{\underline{XX'}} - \lambda(\underline{\underline{U}} - \underline{\underline{YY'}})| = 0,$$

where

$$\underline{\underline{U}} = \underline{\underline{XX'}} + \underline{\underline{YY'}} + \underline{\underline{ZZ'}}.$$

Except for roots that are identically zero, the roots of (2.5) coincide with the nonzero characteristic roots of  $\underline{\underline{X'}}(\underline{\underline{U}} - \underline{\underline{YY'}})^{-1}\underline{\underline{X}}$ . Let

$$(2.6) \quad \underline{\underline{V}} = (\underline{\underline{X}}, \underline{\underline{Y}}, \underline{\underline{U}}),$$

and let

$$(2.7) \quad M(\underline{\underline{V}}) = \underline{\underline{X'}}(\underline{\underline{U}} - \underline{\underline{YY'}})^{-1}\underline{\underline{X}}.$$

The vector of ordered characteristic roots of  $M(\underline{\underline{V}})$  is denoted by

$$(2.8) \quad (\lambda_1, \dots, \lambda_m)' = \underline{\underline{\lambda}}(M(\underline{\underline{V}})),$$

where  $\lambda_1 \geq \dots \geq \lambda_m \geq 0$ . Since the inclusion of zero roots (when  $m > p$ ) causes no trouble in the sequel we assume that the tests depend on  $\underline{\underline{\lambda}}(M(\underline{\underline{V}}))$ .

It is well known that these tests are invariant under certain groups of transformations. See Anderson (1958), sec. 8.10, or Lehmann (1959), sec. 7.9.

The admissibility of these tests can be stated in terms of the geometric characteristics of the acceptance regions. Let

$$(2.9) \quad R_{<}^m = \{\underline{\underline{\lambda}} \in R^m \mid \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0\},$$

$$R_{+}^m = \{\underline{\underline{\lambda}} \in R^m \mid \lambda_1 \geq 0, \dots, \lambda_m \geq 0\}.$$

Definition 2.1. A region  $A$  in  $R_{<}^m$  is monotone if  $\lambda \in A$ ,  $\underline{v} \in R_{<}^m$ ,  $v_i \leq \lambda_i$ ,  $i = 1, \dots, m$ , imply  $\underline{v} \in A$ .

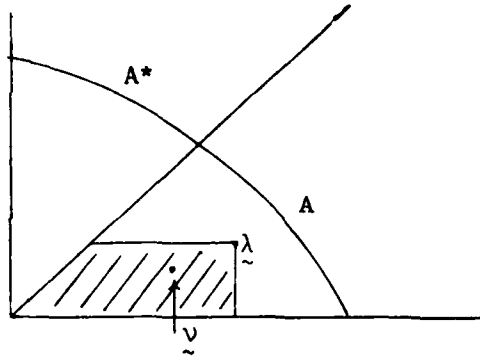


Figure 2.1

Definition 2.2. For  $A \subset R_{<}^m$  the extended region  $A^*$  is defined by

$$A^* = \bigcup_{\pi} \{ (x_{\pi(1)}, \dots, x_{\pi(m)})' \mid \underline{x} \in A \},$$

where  $\pi$  ranges over all permutations of  $(1, \dots, m)$ .

Now we can state the following theorem:

Theorem 2.1. If the region  $A$  in  $R_{<}^m$  is monotone and if the extended region  $A^*$  is closed and convex, then  $A$  is the acceptance region of an admissible test.

This is the essential part of the most important result in the main theorem of Schwartz (1967).

Another characterization of admissible tests is given in terms of majorization.

Definition 2.3. A vector  $\underline{\lambda} = (\lambda_1, \dots, \lambda_m)'$  weakly majorizes a vector  $\underline{v} = (v_1, \dots, v_m)'$  if

$$(2.10) \quad \lambda_{[1]} \geq v_{[1]}, \lambda_{[1]} + \lambda_{[2]} \geq v_{[1]} + v_{[2]}, \dots, \lambda_{[1]} + \dots + \lambda_{[m]} \geq v_{[1]} + \dots + v_{[m]},$$

where  $\lambda_{[i]}$  and  $v_{[i]}$ ,  $i = 1, \dots, m$ , are the coordinates rearranged in nonascending order.

We use the notation

$$\underline{\lambda} \succ_w \underline{v} \text{ or } \underline{v} \prec_w \underline{\lambda}$$

if  $\underline{\lambda}$  weakly majorizes  $\underline{v}$ .

Remark. If  $\underline{\lambda}, \underline{v} \in R^m_{\leq}$  then  $\underline{\lambda} \succ_w \underline{v}$  is simply

$$\lambda_1 \geq v_1, \lambda_1 + \lambda_2 \geq v_1 + v_2, \dots, \lambda_1 + \dots + \lambda_m \geq v_1 + \dots + v_m.$$

If the last inequality in (2.10) is replaced by an equality we say simply that  $\underline{\lambda}$  majorizes  $\underline{v}$  and denote this by

$$(2.11) \quad \underline{\lambda} \succ \underline{v} \text{ or } \underline{v} \prec \underline{\lambda}.$$

The theory of majorization and the related inequalities are developed in detail in Marshall and Olkin (1979).

Definition 2.4. A region  $A$  in  $R^m_{\leq}$  is said to be monotone in majorization if  $\underline{\lambda} \in A$ ,  $\underline{v} \in R^m_{\leq}$ ,  $\underline{v} \prec_w \underline{\lambda}$  imply  $\underline{v} \in A$ .

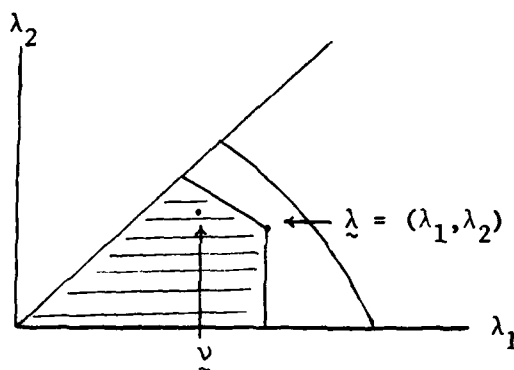


Figure 2.2

Theorem 2.2. If a region  $A$  in  $R_{\leq}^m$  is closed, convex, and monotone in majorization, then  $A$  is the acceptance region of an admissible test.

Theorem 2.1 and Theorem 2.2 are equivalent. It will be seen in Section 4 that Theorem 2.2 can be more conveniently proved. Then an argument about the extreme points of a certain convex set (Lemma 4.12) establishes the equivalence of the two theorems.

Application of the theory of Schur-convex functions yield several corollaries to Theorem 2.2.

Corollary 2.1. Let  $g$  be continuous, nondecreasing, and convex in  $[0,1)$ . Let

$$(2.12) \quad f(\lambda) = f(\lambda_1, \dots, \lambda_m) = \sum_{i=1}^m g(\lambda_i) .$$

Then a test with the acceptance region

$$A = \{\lambda \mid f(\lambda) \leq c\}$$

is admissible.

A proof of this is given in Section 4. Setting  $g(\lambda) = -\log(1-\lambda)$ ,  $g(\lambda) = \lambda/(1-\lambda)$ ,  $g(\lambda) = \lambda$ , respectively, shows that Wilks' likelihood ratio test, the Lawley-Hotelling trace test, and the Bartlett-Nanda-Pillai test are admissible. Admissibility of Roy's maximum root test

$$A: \lambda_1 \leq c$$

can be proved directly from Theorem 2.1 or Theorem 2.2.

We note that (2.12) is a special case of a Schur-convex function. A function  $\phi$  is called Schur-convex if  $\underline{x} \prec \underline{y}$  implies  $\phi(\underline{x}) \leq \phi(\underline{y})$ . The following facts are well known in the theory of Schur-convex functions. (See Marshall and Olkin (1979), Chap. 3.) If  $\phi$  is increasing in each argument and Schur-convex then  $\underline{x} \prec_w \underline{y}$  implies  $\phi(\underline{x}) \leq \phi(\underline{y})$ . If  $\phi$  is symmetric and quasi-convex (that is,  $\{\underline{x} | \phi(\underline{x}) \leq c\}$  is convex for all  $c$ ), then  $\phi$  is Schur-convex. Combining these facts we obtain the following corollary.

Corollary 2.2. If  $f$  is lower semicontinuous, increasing in each argument, symmetric, and quasi-convex, then a test with the acceptance region  $A = \{\underline{\lambda} | f(\underline{\lambda}) \leq c\}$  is admissible.

Corollary 2.2 is identical to the essential part of Theorem 1 of Schwartz (1967). It is interesting to see that he arrived at these

conditions on  $f$  without the notion of Schur-convexity. From our viewpoint the meaning of the conditions are clearer here. Note that Corollary 2.2 is equivalent to Theorem 2.1.

The rest of the paper will be devoted to the proof of the above theorems.

### 3. Stein's Theorem and the Exponential Family.

In this section we state Stein's theorem and show how our problem fits its setting.

An exponential family of distributions consists of a finite-dimensional Euclidean space  $\mathcal{Y}$ , a measure  $m$  on the  $\sigma$ -algebra  $\mathcal{B}$  of all ordinary Borel sets of  $\mathcal{Y}$ , a subset  $\Omega$  of the adjoint space  $\mathcal{Y}'$  (the linear space of all real-valued linear functions on  $\mathcal{Y}$ ) such that

$$(3.1) \quad \psi(\omega) = \int_{\mathcal{Y}} e^{\omega' \mathcal{Y}} dm(\mathcal{Y}) < \infty, \quad \omega \in \Omega,$$

and  $P$ , the function on  $\Omega$  to the set of probability measures on  $\mathcal{B}$  given by

$$P_{\omega}(A) = \frac{1}{\psi(\omega)} \int_A e^{\omega' \mathcal{Y}} dm(\mathcal{Y}), \quad A \in \mathcal{B}.$$

Theorem 3.1 (Stein). Let  $(\mathcal{Y}, m, \Omega, P)$  be an exponential family and  $\Omega_0$  a nonempty proper subset of  $\Omega$ . Let  $A$  be a subset of  $\mathcal{Y}$  such that (i)  $A$  is closed and convex and (ii) for every vector  $\omega \in \mathcal{Y}'$  and real

c for which  $\{\underline{y} | \underline{\omega}'\underline{y} > c\}$  and A are disjoint, there exists  $\underline{\omega}_1 \in \Omega$  such that for arbitrarily large  $\lambda$   $\underline{\omega}_1 + \lambda \underline{\omega} \in \Omega - \Omega_0$ . Then the test with acceptance region A is admissible for testing the hypothesis  $\underline{\omega} \in \Omega_0$  against the alternative  $\underline{\omega} \in \Omega - \Omega_0$ .

Note that A need not be closed if the boundary of A has m-measure zero. (See Proposition 1 in Appendix.)

We rewrite the distribution of  $(\underline{X}, \underline{Y}, \underline{Z})$  in an exponential form. Let  $\underline{U} = \underline{X}\underline{X}' + \underline{Y}\underline{Y}' + \underline{Z}\underline{Z}' = (u_{ij})$  and  $\underline{\Sigma}^{-1} = (\sigma^{ij})$ . For a general matrix  $\underline{C} = (c_1, \dots, c_k)$  let  $\text{Vec}(\underline{C}) = (c_1', \dots, c_k')'$ . The density of  $(\underline{X}, \underline{Y}, \underline{Z})$  can be written as

$$\begin{aligned} (3.2) \quad f(\underline{X}, \underline{Y}, \underline{Z}) &= K(\underline{M}, \underline{N}, \underline{\Sigma}) \exp\{\text{tr } \underline{M}' \underline{\Sigma}^{-1} \underline{X} + \text{tr } \underline{N}' \underline{\Sigma}^{-1} \underline{Y} - \frac{1}{2} \text{tr } \underline{\Sigma}^{-1} \underline{U}\} \\ &= K(\underline{M}, \underline{N}, \underline{\Sigma}) \exp\{\underline{\omega}'_{(1)} \underline{y}_{(1)} + \underline{\omega}'_{(2)} \underline{y}_{(2)} + \underline{\omega}'_{(3)} \underline{y}_{(3)}\}, \end{aligned}$$

where

$$K(\underline{M}, \underline{N}, \underline{\Sigma}) = \frac{\exp\{-\frac{1}{2} \text{tr } \underline{\Sigma}^{-1} (\underline{M}\underline{M}' + \underline{N}\underline{N}')\}}{(2\pi)^{\frac{1}{2} p(m+r+n)} |\underline{\Sigma}|^{\frac{1}{2}(m+r+n)}},$$

$$\begin{aligned} (3.3) \quad \underline{\omega}_{(1)} &= \text{vec}(\underline{\Sigma}^{-1} \underline{M}), \quad \underline{\omega}_{(2)} = \text{vec}(\underline{\Sigma}^{-1} \underline{N}), \\ \underline{\omega}_{(3)} &= -\frac{1}{2} (\sigma^{11}, 2\sigma^{12}, \dots, 2\sigma^{1p}, \sigma^{22}, \dots, \sigma^{pp})', \\ \underline{y}_{(1)} &= \text{vec}(\underline{X}), \quad \underline{y}_{(2)} = \text{vec}(\underline{Y}), \\ \underline{y}_{(3)} &= (u_{11}, u_{12}, \dots, u_{1p}, u_{22}, \dots, u_{pp})'. \end{aligned}$$

If we denote the mapping  $(\underline{X}, \underline{Y}, \underline{Z}) \rightarrow \underline{y} = (\underline{y}'_{(1)}, \underline{y}'_{(2)}, \underline{y}'_{(3)})$  by  $g$ ,

$$(3.4) \quad \underline{y} = g(\underline{X}, \underline{Y}, \underline{Z}),$$

then the measure of a set  $A$  in the space of  $\underline{y}$  is

$$(3.5) \quad m(A) = \mu(g^{-1}(A)),$$

where  $\mu$  is the ordinary Lebesgue measure on  $R^{p(m+r+n)}$ . We note that

$(\underline{X}, \underline{Y}, \underline{U})$  is a sufficient statistic and so is  $\underline{y} = (\underline{y}'_{(1)}, \underline{y}'_{(2)}, \underline{y}'_{(3)})'$ .

Because a test which is admissible with respect to the class of tests based on a sufficient statistic is admissible in the whole class of tests, we consider only tests based on a sufficient statistic. Then the acceptance regions of these tests are subsets in the space of  $\underline{y}$ . The density of  $\underline{y}$  given by the right hand side of (3.2) is of the form of the exponential family and therefore we can apply Stein's theorem. Furthermore, since the transformation  $(\underline{X}, \underline{Y}, \underline{U}) \rightarrow \underline{y}$  is linear, we prove the convexity of an acceptance region in terms of  $(\underline{X}, \underline{Y}, \underline{U})$ . The acceptance region of an invariant test is given in terms of  $\underline{\lambda}(M(\underline{V})) = (\lambda_1, \dots, \lambda_m)'$ . Therefore, in order to prove the admissibility of these tests we have to check that the inverse image of  $A$

$$\{\underline{V} \mid \underline{\lambda}(M(\underline{V})) \in A\}$$

satisfies the conditions of Stein's theorem. We carry this out in the next section.

For the sequel we use the following notation:

$\underline{A} \geq \underline{B}$  if and only if  $\underline{A} - \underline{B}$  is positive semidefinite,

$\underline{A} > \underline{B}$  if and only if  $\underline{A} - \underline{B}$  is positive definite.

#### 4. Proofs.

We start with lemmas concerning matrix inequalities.

##### Lemma 4.1.

$$(4.1) \quad p\tilde{U}_1 + q\tilde{U}_2 - (p\tilde{Y}_1 + q\tilde{Y}_2)(p\tilde{Y}_1 + q\tilde{Y}_2)' \\ \geq p(\tilde{U}_1 - \tilde{Y}_1\tilde{Y}_1') + q(\tilde{U}_2 - \tilde{Y}_2\tilde{Y}_2') .$$

Proof. The left-hand side minus the right-hand side is

$$(4.2) \quad p\tilde{Y}_1\tilde{Y}_1' + q\tilde{Y}_2\tilde{Y}_2' - p^2\tilde{Y}_1\tilde{Y}_1' - q^2\tilde{Y}_2\tilde{Y}_2' - pq(\tilde{Y}_1\tilde{Y}_2' + \tilde{Y}_2\tilde{Y}_1') \\ = p(1-p)\tilde{Y}_1\tilde{Y}_1' + q(1-q)\tilde{Y}_2\tilde{Y}_2' - pq(\tilde{Y}_1\tilde{Y}_2' + \tilde{Y}_2\tilde{Y}_1') \\ = pq(\tilde{Y}_1 - \tilde{Y}_2)(\tilde{Y}_1 - \tilde{Y}_2)' \geq 0 . \quad \text{Q.E.D.}$$

Lemma 4.2. If  $\tilde{A} \geq \tilde{B} > 0$ , then  $\tilde{A}^{-1} \leq \tilde{B}^{-1}$ .

Proof. Let  $\tilde{A} = \tilde{F}\tilde{D}\tilde{F}'$  and  $\tilde{B} = \tilde{F}\tilde{F}'$  where  $\tilde{D}$  is diagonal and  $\tilde{F}$  is nonsingular. Then  $\tilde{A} \geq \tilde{B}$  implies  $\tilde{D} \geq \tilde{I}$ , and  $\tilde{B}^{-1} - \tilde{A}^{-1} = (\tilde{F}')^{-1}(\tilde{I} - \tilde{D}^{-1})\tilde{F}^{-1} \geq 0$  because  $\tilde{I} - \tilde{D}^{-1} \geq 0$ . Q.E.D.

Lemma 4.3. If  $\tilde{A} > 0$ , then  $f(\tilde{x}, \tilde{A}) = \tilde{x}'\tilde{A}^{-1}\tilde{x}$  is convex in  $(\tilde{x}, \tilde{A})$ .

Proof. If  $\tilde{A} = \tilde{D}$  is diagonal, then

$$f(\tilde{x}, \tilde{D}) = \tilde{x}'\tilde{D}^{-1}\tilde{x} = \sum_{i=1}^p x_i^2 / d_{ii}$$

is a convex function because it is the sum of convex functions  $x_i^2/d_{ii}$ .

(The matrix of second-order partial derivatives of  $x^2/d$  is positive semidefinite.) In general, we can write

$$\tilde{A}_1 = \tilde{F} \tilde{D} \tilde{F}' , \quad \tilde{A}_2 = \tilde{F} \tilde{F}' .$$

Hence

$$\begin{aligned} (4.3) \quad & (p\tilde{x}_1 + q\tilde{x}_2)' (p\tilde{A}_1 + q\tilde{A}_2)^{-1} (p\tilde{x}_1 + q\tilde{x}_2) \\ &= (p\tilde{F}^{-1}\tilde{x}_1 + q\tilde{F}^{-1}\tilde{x}_2)' (p\tilde{D} + q\tilde{I})^{-1} (p\tilde{F}^{-1}\tilde{x}_1 + q\tilde{F}^{-1}\tilde{x}_2) \\ &\leq p(\tilde{F}^{-1}\tilde{x}_1)' \tilde{D}^{-1} (\tilde{F}^{-1}\tilde{x}_1) + q(\tilde{F}^{-1}\tilde{x}_2)' (\tilde{F}^{-1}\tilde{x}_2) \\ &= p\tilde{x}_1' \tilde{A}_1^{-1} \tilde{x}_1 + q\tilde{x}_2' \tilde{A}_2^{-1} \tilde{x}_2 . \end{aligned} \quad \text{Q.E.D.}$$

Lemma 4.4. If  $\tilde{A}_1 > 0$ ,  $\tilde{A}_2 > 0$ , then

$$\begin{aligned} (4.4) \quad & (p\tilde{B}_1 + q\tilde{B}_2)' (p\tilde{A}_1 + q\tilde{A}_2)^{-1} (p\tilde{B}_1 + q\tilde{B}_2) \\ &\leq p\tilde{B}_1' \tilde{A}_1^{-1} \tilde{B}_1 + q\tilde{B}_2' \tilde{A}_2^{-1} \tilde{B}_2 . \end{aligned}$$

Proof. From Lemma 4.4, we have for all  $\tilde{y}$

$$\begin{aligned} & p\tilde{y}' \tilde{B}_1' \tilde{A}_1^{-1} \tilde{B}_1 \tilde{y} + q\tilde{y}' \tilde{B}_2' \tilde{A}_2^{-1} \tilde{B}_2 \tilde{y} - \tilde{y}' (p\tilde{B}_1 + q\tilde{B}_2)' (p\tilde{A}_1 + q\tilde{A}_2)^{-1} (p\tilde{B}_1 + q\tilde{B}_2) \tilde{y} \\ &= p(\tilde{B}_1 \tilde{y})' \tilde{A}_1^{-1} (\tilde{B}_1 \tilde{y}) + q(\tilde{B}_2 \tilde{y})' \tilde{A}_2^{-1} (\tilde{B}_2 \tilde{y}) \\ &\quad - (p\tilde{B}_1 \tilde{y} + q\tilde{B}_2 \tilde{y})' (p\tilde{A}_1 + q\tilde{A}_2)^{-1} (p\tilde{B}_1 \tilde{y} + q\tilde{B}_2 \tilde{y}) \\ &\geq 0 . \end{aligned}$$

Thus the matrix of the quadratic form in  $y$  is positive semidefinite. Q.E.D.

The relation as in (4.4) is sometimes called "matrix convexity". (See Marshall and Olkin (1979).)

Theorem 4.1.

$$M(pV_1 + qV_2) \leq pM(V_1) + qM(V_2),$$

where

$$V_1 = (X_1, Y_1, U_1), \quad V_2 = (X_2, Y_2, U_2),$$

$$U_1 - Y_1 Y_1' > 0, \quad U_2 - Y_2 Y_2' > 0,$$

$$0 \leq p = 1 - q \leq 1.$$

Proof. Lemma 4.1 and Lemma 4.2 show that

$$(4.5) \quad [pU_1 + qU_2 - (pY_1 + qY_2)(pY_1 + qY_2)']^{-1} \\ \leq [p(U_1 - Y_1 Y_1') + q(U_2 - Y_2 Y_2')]^{-1}.$$

This implies

$$(4.6) \quad M(pV_1 + qV_2) \leq (pX_1 + qX_2)' [p(U_1 - Y_1 Y_1') + q(U_2 - Y_2 Y_2')]^{-1} (pX_1 + qX_2).$$

Then Lemma 4.4 implies that the right-hand side of (4.6) is less than or equal to

$$pX_1'(U_1 - Y_1 Y_1')^{-1} X_1 + qX_2'(U_2 - Y_2 Y_2')^{-1} X_2 = pM(V_1) + qM(V_2) . \quad \text{Q.E.D.}$$

For the next step we need several lemmas concerning majorization. Proofs are given here so that the paper is self-contained. For a full discussion see Marshall and Olkin (1979).

Lemma 4.5.

$$(4.7) \quad \sum_{i=1}^k \lambda_i(A) = \max_{\tilde{R}'\tilde{R} = I_k} \text{tr } \tilde{R}'A\tilde{R} .$$

Proof.

$$(4.8) \quad \max_{\tilde{R}'\tilde{R} = I_k} \text{tr } \tilde{R}'A\tilde{R} = \max_{\tilde{Q}'\tilde{Q} = I_k} \text{tr } \tilde{Q}'D\tilde{Q} ,$$

where  $\tilde{D}$  is a diagonal matrix with

$$(d_{11}, \dots, d_{mm})' = \lambda(\tilde{A}) .$$

If  $\tilde{Q}$  is the appropriate submatrix of the identity matrix, equality in (4.7) is achieved. To prove the inequality we augment  $\tilde{Q}$  to a full orthogonal matrix  $\tilde{G}$  and let

$$\tilde{B} = \tilde{G}'\tilde{D}\tilde{G} = (b_{ij}) .$$

Then

$$\text{tr } \tilde{Q}'\tilde{D}\tilde{Q} = b_{11} + \dots + b_{kk} .$$

Note that

$$b_{jj} = \sum_{i=1}^m \lambda_i g_{ij}^2$$

or

$$(b_{11}, \dots, b_{mm})' = P\lambda,$$

where

$$P = (p_{ij}) = (g_{ij}^2).$$

$P$  is a doubly stochastic matrix. Then (4.7) follows from  $\lambda \succ_w P\lambda$  (Lemma 4.6). Q.E.D.

Lemma 4.6. If  $P$  is an  $m \times m$  doubly stochastic matrix, then

$$y \succ_w Py.$$

Proof.

$$(4.9) \quad \sum_{i=1}^k \sum_{j=1}^m p_{ij} y_j = \sum_{j=1}^m g_j y_j,$$

where

$$g_j = \sum_{i=1}^k p_{ij} \quad (0 \leq g_j \leq 1, \quad \sum_{j=1}^m g_j = k).$$

Then

$$\sum_{j=1}^m g_j y_j - \sum_{i=1}^k y_i = \sum_{i=1}^m g_i y_i - \sum_{i=1}^k y_i + y_k (k - \sum_{i=1}^m g_i)$$

$$= \sum_{i=1}^k (y_i - y_k)(g_i - 1) + \sum_{i=k+1}^m (y_i - y_k)g_i$$

$\leq 0$ .

Q.E.D.

Remark. Lemma 4.6 holds with  $\prec$  replacing  $\prec_w$ .

Lemma 4.7. If  $A \preceq B$ , then  $\lambda(A) \prec_w \lambda(B)$ .

Proof. From Lemma 4.5

$$(4.10) \quad \sum_{i=1}^k \lambda_i(A) = \max_{R'R = I_k} \text{tr } R'AR \leq \max_{R'R = I_k} \text{tr } R'BR = \sum_{i=1}^k \lambda_i(B).$$

Remark. Actually  $A \preceq B$  implies  $\lambda_i(A) \leq \lambda_i(B)$ ,  $i = 1, \dots, m$ , but this stronger result is not needed.

Lemma 4.8.

$$\lambda(A+B) \prec_w \lambda(A) + \lambda(B).$$

Proof.

$$\begin{aligned} (4.11) \quad \sum_{i=1}^k \lambda_i(A+B) &= \max_{R'R = I_k} \text{tr } R'(A+B)R \\ &\leq \max_{R'R = I_k} \text{tr } R'AR + \max_{R'R = I_k} \text{tr } R'BR \\ &= \sum_{i=1}^k \lambda_i(A) + \sum_{i=1}^k \lambda_i(B) \\ &= \sum_{i=1}^k \{\lambda_i(A) + \lambda_i(B)\}. \end{aligned}$$

Q.E.D.

Remark. Lemma 4.8 holds with  $\prec$  replacing  $\prec_w$ .

Now the matrix inequality of Theorem 4.1 translates into majorization of vectors of characteristic roots.

Theorem 4.2.

$$\lambda(M(pV_1 + qV_2)) \prec_w p\lambda(M(V_1)) + q\lambda(M(V_2)).$$

Proof. Theorem 4.1 and Lemma 4.7 imply

$$(4.12) \quad \lambda(M(pV_1 + qV_2)) \prec_w \lambda(pM(V_1) + qM(V_2)).$$

Then by Lemma 4.8

$$(4.13) \quad \lambda(pM(V_1) + qM(V_2)) \prec_w p\lambda(M(V_1)) + q\lambda(M(V_2)).$$

From (4.12) and (4.13) Theorem 4.2 follows. Q.E.D.

Now let  $A$  be a region in the space of roots and let  $\tilde{A}$  be the inverse image of  $A$  in  $V (= (X, Y, U))$  space,

$$\tilde{A} = \{V \mid \lambda(M(V)) \in A\}.$$

We want to show that  $\tilde{A}$  is convex for a region satisfying the condition of Theorem 2.2. Let  $V_i = (X_i, Y_i, U_i) \in \tilde{A}$ ,  $i = 1, 2$ . Then  $\lambda(M(V_i)) \in A$ ,  $i = 1, 2$ , and by convexity of  $A$  we have  $p\lambda(M(V_1)) + q\lambda(M(V_2)) \in A$ . Then by Theorem 4.2 and monotonicity in majorization of  $A$

$$\lambda(M(pV_1 + qV_2)) \in A.$$

Hence  $p\tilde{V}_1 + q\tilde{V}_2 \in \tilde{A}$  and  $\tilde{A}$  is convex. Furthermore the boundary of  $\tilde{A}$  has probability 0. (See Propositions 1 and 2 in Appendix.) Therefore, condition (i) of Stein's theorem is satisfied.

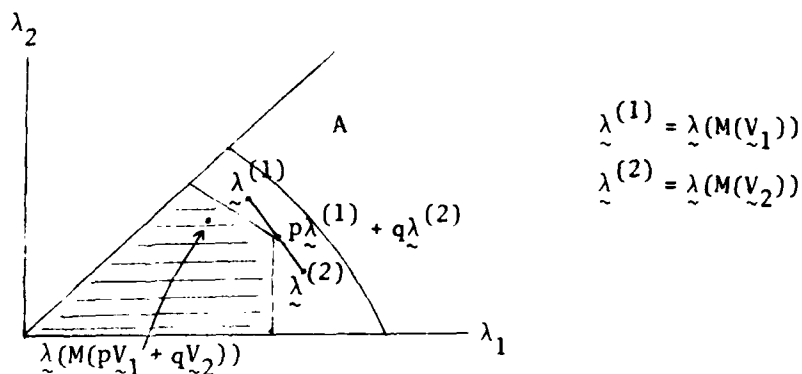


Figure 4.1

For the condition (ii) we have the following lemma. The proof was suggested by Charles Stein.

Lemma 4.9. For the acceptance region  $A$  of Theorem 2.1 or Theorem 2.2 the condition (ii) of Stein's theorem is satisfied.

Proof. Let  $\omega$  correspond to  $(\phi, \psi, \theta)$  then

$$\begin{aligned}
 (4.14) \quad \omega'_Y &= \omega'_{(1)Y(1)} + \omega'_{(2)Y(2)} + \omega'_{(3)Y(3)} \\
 &= \text{tr } \phi'X + \text{tr } \psi'Y - \frac{1}{2} \text{tr } \theta U,
 \end{aligned}$$

where  $\tilde{\Theta}$  is symmetric. Suppose that  $\{y|\omega'y > c\}$  is disjoint from  $\tilde{A}$ . We want to show that in this case  $\tilde{\Theta}$  is positive semidefinite. If this were not true, then

$$\tilde{\Theta} = \tilde{F} \begin{bmatrix} -\tilde{I} & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{I} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} \end{bmatrix} \tilde{F}',$$

where  $\tilde{F}$  is nonsingular and  $-\tilde{I}$  is not vacuous. Let

$$\tilde{X} = (1/\gamma)\tilde{X}_0, \quad \tilde{Y} = (1/\gamma)\tilde{Y}_0,$$

$$\tilde{U} = (\tilde{F}')^{-1} \begin{bmatrix} \gamma\tilde{I} & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{I} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{I} \end{bmatrix} \tilde{F}^{-1},$$

$$\tilde{V} = (\tilde{X}, \tilde{Y}, \tilde{U}),$$

where  $\tilde{X}_0, \tilde{Y}_0$  are fixed matrices and  $\gamma$  is a positive number. Then

$$(4.15) \quad \omega'y = \frac{1}{\gamma} \text{tr } \tilde{\Phi}'\tilde{X}_0 + \frac{1}{\gamma} \text{tr } \tilde{\Psi}'\tilde{Y}_0 + \frac{1}{2} \text{tr} \begin{bmatrix} \gamma\tilde{I} & \tilde{0} & \tilde{0} \\ \tilde{0} & -\tilde{I} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} \end{bmatrix} > c$$

for sufficiently large  $\gamma$ . On the other hand,

$$(4.16) \quad \lambda(\tilde{M}(\tilde{V})) = \lambda\{\tilde{X}'(\tilde{U}-\tilde{Y}\tilde{Y}')^{-1}\tilde{X}\}$$

$$= \frac{1}{\gamma^2} \lambda \left\{ \tilde{X}_0' (\tilde{F}')^{-1} \begin{bmatrix} \gamma\tilde{I} & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{I} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{I} \end{bmatrix} \tilde{F}^{-1} - \frac{1}{\gamma^2} \tilde{Y}_0 \tilde{Y}_0' \right\}^{-1} \tilde{X}_0$$

$$\rightarrow 0 \quad \text{as } \gamma \rightarrow \infty.$$

Therefore,  $\underline{v} \in \tilde{A}$  for sufficiently large  $\gamma$ . This is a contradiction.

Hence  $\underline{\Theta}$  is positive semidefinite.

Now let  $\underline{\omega}_1$  correspond to  $(\underline{\phi}_1, \underline{0}, \underline{I})$ , where  $\underline{\phi}_1 \neq \underline{0}$ . Then  $\underline{I} + \lambda \underline{\Theta}$  is positive definite and  $\underline{\phi}_1 + \lambda \underline{\phi} \neq \underline{0}$  for sufficiently large  $\lambda$ . Hence,

$$\underline{\omega}_1 + \lambda \underline{\omega} \in \Omega - \Omega_0$$

for sufficiently large  $\lambda$ . Q.E.D.

Now we have proved Theorem 2.2.

Proof of Corollary 2.1. Being a sum of convex functions  $f$  is convex and hence  $A$  is convex.  $A$  is closed because  $f$  is continuous. We want to show that if  $f(\underline{x}) \leq c$  and  $\underline{y} <_w \underline{x}$  ( $\underline{x}, \underline{y} \in R^m$ ) then  $f(\underline{y}) \leq c$ . Let  $\tilde{x}_k = \sum_{i=1}^k x_i$ ,  $\tilde{y}_k = \sum_{i=1}^k y_i$ . Then  $\underline{y} <_w \underline{x}$  if and only if  $\tilde{x}_k \geq \tilde{y}_k$ ,  $k = 1, \dots, m$ . Let  $f(\underline{x}) = h(\tilde{x}_1, \dots, \tilde{x}_m) = g(\tilde{x}_1) + \sum_{i=2}^m g(\tilde{x}_i - \tilde{x}_{i-1})$ . It suffices to show that  $h(\tilde{x}_1, \dots, \tilde{x}_m)$  is increasing in each  $\tilde{x}_i$ . For  $i \leq m-1$  the convexity of  $g$  implies that

$$\begin{aligned} h(\tilde{x}_1, \dots, \tilde{x}_i + \epsilon, \dots, \tilde{x}_m) - h(\tilde{x}_1, \dots, \tilde{x}_i, \dots, \tilde{x}_m) \\ = g(\tilde{x}_i + \epsilon) - g(\tilde{x}_i) - \{g(\tilde{x}_{i+1}) - g(\tilde{x}_{i+1} - \epsilon)\} \\ \geq 0. \end{aligned}$$

For  $i = m$  the monotonicity of  $g$  implies

$$h(\tilde{x}_1, \dots, \tilde{x}_m + \epsilon) - h(\tilde{x}_1, \dots, \tilde{x}_m) = g(\tilde{x}_m + \epsilon) - g(\tilde{x}_m) \geq 0. \text{ Q.E.D.}$$

Lemma 4.10.  $A \subset R_{<}^m$  is convex and monotone in majorization if and only if  $A$  is monotone and  $A^*$  is convex.

Proof.

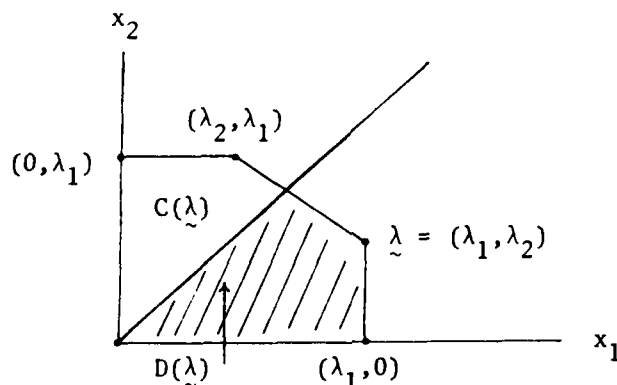
Necessity. If  $A$  is monotone in majorization, then it is obviously monotone. By Proposition 3 of Appendix  $A^*$  is convex.

Sufficiency. For  $\lambda \in R_{<}^m$  let

$$(4.17) \quad C(\lambda) = \{x \mid x \in R_{+}^m, x \leq_w \lambda\},$$

$$D(\lambda) = \{x \mid x \in R_{<}^m, x \leq_w \lambda\}.$$

It will be proved in Lemma 4.12 and its corollary that monotonicity of  $A$  and convexity of  $A^*$  implies  $C(\lambda) \subset A^*$ . Then  $D(\lambda) = C(\lambda) \cap R_{<}^m \subset A^* \cap R_{<}^m = A$ . Now suppose  $v \in R_{<}^m$  and  $v \leq_w \lambda$ . Then  $v \in D(\lambda) \subset A$ . This shows that  $A$  is monotone in majorization. Furthermore, if  $A^*$  is convex, then  $A = R_{<}^m \cap A^*$  is convex. Q.E.D.



• = extreme points

Figure 4.2

Lemma 4.11. Let  $C$  and  $D$  be closed and convex. If the extreme points of  $C$  are contained in  $D$ , then  $C \subset D$ .

Proof. Obvious.

Now we present the following key lemma.

Lemma 4.12. The extreme points of  $C(\lambda)$  are all vectors of the form

$$(4.18) \quad (\delta_{\pi(1)}\lambda_{\pi(1)}, \dots, \delta_{\pi(m)}\lambda_{\pi(m)}) ,$$

where  $\pi$  is a permutation of  $(1, \dots, m)$  and  $\delta_1 = \dots = \delta_k = 1$ ,  
 $\delta_{k+1} = \dots = \delta_m = 0$  for some  $k$ .

Proof.  $C(\lambda)$  is convex. (See Proposition 4 of Appendix). Now note that  $C(\lambda)$  is permutation symmetric, that is, if  $(x_1, \dots, x_m)' \in C(\lambda)$ , then  $(x_{\pi(1)}, \dots, x_{\pi(m)})' \in C(\lambda)$  for any permutation  $\pi$ . Therefore, for any permutation  $\pi$ ,  $\pi(C(\lambda)) = \{(x_{\pi(1)}, \dots, x_{\pi(m)})' \mid x \in C(\lambda)\}$  coincides with  $C(\lambda)$ . This implies that if  $(x_1, \dots, x_m)'$  is an extreme point of  $C(\lambda)$ , then  $(x_{\pi(1)}, \dots, x_{\pi(m)})'$  is also an extreme point. In particular,  $(x_{[1]}, \dots, x_{[m]})' \in R_{\leq}^m$  is an extreme point. Conversely if  $(x_1, \dots, x_m) \in R_{\leq}^m$  is an extreme point of  $C(\lambda)$ , then  $(x_{\pi(1)}, \dots, x_{\pi(m)})'$  is an extreme point.

We see that once we enumerate the extreme points of  $C(\lambda)$  in  $R_{\leq}^m$ , the rest of the extreme points can be obtained by permutation.

Suppose  $x \in R_{\leq}^m$ . An extreme point being the intersection of  $m$  hyperplanes has to satisfy  $m$  of the following  $2m$  equations:

$$(4.19) \quad \begin{array}{ll} E_1: x_1 = 0, & F_1: x_1 = \lambda_1, \\ E_2: x_2 = 0, & F_2: x_1 + x_2 = \lambda_1 + \lambda_2, \\ \vdots & \vdots \\ E_m: x_m = 0, & F_m: x_1 + \dots + x_m = \lambda_1 + \dots + \lambda_m. \end{array}$$

Suppose that  $k$  is the first index such that  $E_k$  holds (namely, suppose that  $F_1, \dots, F_{k-1}, E_k$  hold). Then  $\underline{x} \in R_{\leq}^m$  implies

$$0 = x_k \geq x_{k+1} \geq \dots \geq x_m = 0.$$

Therefore,  $E_{k+1}, \dots, E_m$  hold. This gives the point  $\underline{\lambda} = (\lambda_1, \dots, \lambda_{k-1}, 0, \dots, 0)$  which is in  $R_{\leq}^m \cap C(\underline{\lambda})$ . Therefore  $\underline{\lambda}$  is an extreme point. Q.E.D.

Remark. If  $\lambda_1 > \lambda_2 > \dots > \lambda_m$  above, the total number of distinct extreme points is

$$\sum_{k=0}^m \frac{m!}{(m-k)!} = m! \sum_{k=0}^m \frac{1}{k!} \doteq m! e.$$

This follows from the fact that by permutation of  $\underline{\lambda} = (\lambda_1, \dots, \lambda_k, 0, \dots, 0)'$  we obtain  $m!/(m-k)!$  distinct points.

Mirsky (1959) gave the explicit expression of the extreme points in the form

$$(4.20) \quad (\delta_1 \lambda_{\pi(1)}, \dots, \delta_m \lambda_{\pi(m)})',$$

where  $\pi$  is a permutation and  $\delta_i$ 's are zero or one. Actually this set of points include some points that are not extreme points.

#### Corollary 4.1.

$$C(\underline{\lambda}) \subset A^*.$$

Proof. If  $A$  is monotone, then  $A^*$  is monotone in the sense that if  $\underline{\lambda} = (\lambda_1, \dots, \lambda_m)' \in A^*$ ,  $\underline{v} = (v_1, \dots, v_m)'$ ,  $v_i \leq \lambda_i$ ,  $i = 1, \dots, m$ ,

then  $\underline{v} \in A^*$ . (See Proposition 5 of Appendix.) Now the extreme points of  $C(\underline{\lambda})$  given by (4.18) are in  $A^*$  because of permutation symmetry and monotonicity of  $A^*$ . Hence by Lemma 4.11  $C(\underline{\lambda}) \subset A^*$ . Q.E.D.

Proof of Theorem 4.1. Immediate from Theorem 4.2 and Lemma 4.10. Q.E.D.

Appendix.

We prove here some technical propositions to complete the argument in Section 4.

Proposition 1. The condition (i) in Stein's theorem can be replaced by (i')  $A$  is convex and the boundary of  $A$  has  $m$ -measure zero.

Proof. If  $A$  is convex, then  $\bar{A}$  (= closure of  $A$ ) is convex. Furthermore,

$$\begin{aligned} A \cap \{y \mid \omega'y > c\} = \emptyset &\Rightarrow A \subset \{y \mid \omega'y \leq c\} \\ &\Rightarrow \bar{A} \subset \{y \mid \omega'y \leq c\} \Rightarrow \bar{A} \cap \{y \mid \omega'y > c\} = \emptyset \end{aligned}$$

Therefore Stein's theorem holds with  $A$  replaced by  $\bar{A}$ . Finally,  $m(\bar{A}-A) = 0$  implies that tests with the acceptance regions  $A$  and  $\bar{A}$  are equivalent (for any  $\omega \in \Omega$ ).

Proposition 2. The boundary of  $\tilde{A}$  in the proof of Theorem 4.11 has  $m$ -measure zero.

Proof. We claim that

$$\text{closure of } \tilde{A} \subset \tilde{A} \cup \{V \mid U-YY' \text{ is singular}\} = \tilde{A} \cup C,$$

where  $C = \{V \mid U-YY' \text{ is singular}\}$ . Obviously  $m(C) = 0$ . This implies  $m(C) \supset m(\text{boundary of } \tilde{A}) = 0$ . Now suppose  $V = (X, Y, U) \in \text{closure of } \tilde{A}$ .

Then  $\underline{V} = \lim \underline{V}_i = \lim(\underline{X}_i, \underline{Y}_i, \underline{U}_i)$ , where  $\underline{V}_i \in \tilde{A}$  or  $\lambda(M(\underline{V}_i)) \in A$ . If  $\underline{U} - \underline{Y}\underline{Y}'$  is singular,  $\underline{V} \in C$ . If  $\underline{U} - \underline{Y}\underline{Y}'$  is nonsingular,

$$\begin{aligned} M(\underline{V}) &= \underline{X}'(\underline{U} - \underline{Y}\underline{Y}')^{-1}\underline{X} \\ &= \lim M(\underline{V}_i) = \lim \underline{X}_i'(\underline{U}_i - \underline{Y}_i \underline{Y}_i')^{-1} \underline{X}_i, \end{aligned}$$

$$\lambda(M(\underline{V})) = \lim \lambda(M(\underline{V}_i))$$

by continuity. Since  $A$  is closed,  $\lambda(M(\underline{V})) \in A$ . Then  $\underline{V} \in \tilde{A}$ . Q.E.D.

Proposition 3. If  $A \subset R_{<}^m$  is convex and monotone in majorization, then  $A^*$  is convex.

Proof. Suppose  $\underline{x}, \underline{y} \in A^*$ . For a vector  $\underline{z}$  let  $\underline{z}_{\downarrow}$  denote  $\underline{z}_{\downarrow} = (z_{[1]}, \dots, z_{[m]})' \in R_{<}^m$ . Now

$$(\underline{px} + \underline{qy})_{\downarrow} \prec_w \underline{px}_{\downarrow} + \underline{qy}_{\downarrow}$$

because

$$\max_i (\underline{px}_i + \underline{qy}_i) \leq p \max_i x_i + q \max_i y_i$$

$$\max_{(i,j)} \{(\underline{px}_i + \underline{qy}_i) + (\underline{px}_j + \underline{qy}_j)\}$$

$$\leq p \max_{(i,j)} (x_i + x_j) + q \max_{(i,j)} (y_i + y_j),$$

etc. Hence,

$$(\underline{px} + \underline{qy})_{\downarrow} \in A \quad \text{and} \quad \underline{px} + \underline{qy} \in A^* . \quad \text{Q.E.D.}$$

Proposition 4.  $C(\lambda)$  is convex.

Proof. Let  $\underline{x}, \underline{y} \prec_w \lambda$ . Then  $\underline{x}_\downarrow, \underline{y}_\downarrow \prec_w \lambda$ . As in the proof of Proposition 3 we have

$$(\underline{px} + \underline{qy})_\downarrow \prec_w \underline{px}_\downarrow + \underline{qy}_\downarrow.$$

Now  $\underline{px}_\downarrow + \underline{qy}_\downarrow \prec_w \underline{p}\lambda + \underline{q}\lambda = \lambda$ . Hence

$$(\underline{px} + \underline{qy})_\downarrow \prec_w \lambda,$$

and

$$\underline{px} + \underline{qy} \prec_w \lambda.$$

It is obvious that if  $\underline{x}, \underline{y} \in R_+^m$ , then  $\underline{px} + \underline{qy} \in R_+^m$ . Q.E.D.

Proposition 5. If  $A$  is monotone, then  $A^*$  is monotone.

Proof. Let  $\underline{y} \in A^*$  and for  $\underline{x}$  assume  $x_i \leq y_i$ ,  $i = 1, \dots, m$ . The point of this proposition is that the permutations transforming  $\underline{x} \rightarrow \underline{x}_\downarrow$ ,  $\underline{y} \rightarrow \underline{y}_\downarrow$  might be different. Note the relation

$$x_{[k]} = \max_{(i_1, \dots, i_k)} \{\min(x_{i_1}, \dots, x_{i_k})\}.$$

Then for any  $(i_1, \dots, i_k)$

$$\begin{aligned} \min(x_{i_1}, \dots, x_{i_k}) &\leq \min(y_{i_1}, \dots, y_{i_k}) \\ &\leq \max_{(j_1, \dots, j_k)} \{\min(y_{j_1}, \dots, y_{j_k})\} = y_{[k]}. \end{aligned}$$

Hence  $x_{[k]} \leq y_{[k]}$ ,  $k = 1, \dots, m$ . This with  $y_{\downarrow} \in A$  implies  $x_{\downarrow} \in A$ .

Hence  $\tilde{x} \in A^*$ .

#### Acknowledgment

We are grateful to Ingram Olkin and Charles Stein for their valuable suggestions.

### References

- Anderson, T. W. (1958), An Introduction to Multivariate Statistical Analysis. Wiley, New York.
- Ghosh, M. N. (1964), On the admissibility of some tests of MANOVA, Ann. Math. Statist., 35, 789-794.
- Lehmann, E. L. (1959), Testing Statistical Hypotheses. Wiley, New York.
- Marshall, A. W. and Olkin, I. (1979), Inequalities: Theory of Majorization and Its Applications. Academic Press, New York.
- Mirsky, L. (1959), On a convex set of matrices, Arch. Math., 10, 88-92.
- Schwartz, R. (1967), Admissible tests in multivariate analysis of variance, Ann. Math. Statist., 38, 698-710.
- Stein, C. (1956), The admissibility of Hotelling's  $T^2$ -test, Ann. Math. Statist., 27, 616-623.

# TECHNICAL REPORTS

OFFICE OF NAVAL RESEARCH CONTRACT N00014-67-A-0112-0030 (NR-042-034)

1. "Confidence Limits for the Expected Value of an Arbitrary Bounded Random Variable with a Continuous Distribution Function," T. W. Anderson, October 1, 1969.
2. "Efficient Estimation of Regression Coefficients in Time Series," T. W. Anderson, October 1, 1970.
3. "Determining the Appropriate Sample Size for Confidence Limits for a Proportion," T. W. Anderson and H. Burstein, October 15, 1970.
4. "Some General Results on Time-Ordered Classification," D. V. Hinkley, July 30, 1971.
5. "Tests for Randomness of Directions against Equatorial and Bimodal Alternatives," T. W. Anderson and M. A. Stephens, August 30, 1971.
6. "Estimation of Covariance Matrices with Linear Structure and Moving Average Processes of Finite Order," T. W. Anderson, October 29, 1971.
7. "The Stationarity of an Estimated Autoregressive Process," T. W. Anderson, November 15, 1971.
8. "On the Inverse of Some Covariance Matrices of Toeplitz Type," Raul Pedro Mentz, July 12, 1972.
9. "An Asymptotic Expansion of the Distribution of "Studentized" Classification Statistics," T. W. Anderson, September 10, 1972.
10. "Asymptotic Evaluation of the Probabilities of Misclassification by Linear Discriminant Functions," T. W. Anderson, September 28, 1972.
11. "Population Mixing Models and Clustering Algorithms," Stanley L. Sclove, February 1, 1973.
12. "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance," John James Miller, November 21, 1973.
13. "Maximum Likelihood Estimation in the Birth-and-Death Process," Niels Keiding, November 28, 1973.
14. "Random Orthogonal Set Functions and Stochastic Models for the Gravity Potential of the Earth," Steffen L. Lauritzen, December 27, 1973.
15. "Maximum Likelihood Estimation of Parameters of an Autoregressive Process with Moving Average Residuals and Other Covariance Matrices with Linear Structure," T. W. Anderson, December, 1973.
16. "Note on a Case-Study in Box-Jenkins Seasonal Forecasting of Time series," Steffen L. Lauritzen, April, 1974.

TECHNICAL REPORTS (continued)

17. "General Exponential Models for Discrete Observations,"  
Steffen L. Lauritzen, May, 1974.
18. "On the Interrelationships among Sufficiency, Total Sufficiency and  
Some Related Concepts," Steffen L. Lauritzen, June, 1974.
19. "Statistical Inference for Multiply Truncated Power Series Distributions,"  
T. Cacoullos, September 30, 1974.

Office of Naval Research Contract N00014-75-C-0442 (NR-042-034)

20. "Estimation by Maximum Likelihood in Autoregressive Moving Average Models  
in the Time and Frequency Domains," T. W. Anderson, June 1975.
21. "Asymptotic Properties of Some Estimators in Moving Average Models,"  
Raul Pedro Mentz, September 8, 1975.
22. "On a Spectral Estimate Obtained by an Autoregressive Model Fitting,"  
Mituaki Huzii, February 1976.
23. "Estimating Means when Some Observations are Classified by Linear  
Discriminant Function," Chien-Pai Han, April 1976.
24. "Panels and Time Series Analysis: Markov Chains and Autoregressive  
Processes," T. W. Anderson, July 1976.
25. "Repeated Measurements on Autoregressive Processes," T. W. Anderson,  
September 1976.
26. "The Recurrence Classification of Risk and Storage Processes,"  
J. Michael Harrison and Sidney I. Resnick, September 1976.
27. "The Generalized Variance of a Stationary Autoregressive Process,"  
T. W. Anderson and Raul P. Mentz, October 1976.
28. "Estimation of the Parameters of Finite Location and Scale Mixtures,"  
Javad Behboodian, October 1976.
29. "Identification of Parameters by the Distribution of a Maximum  
Random Variable," T. W. Anderson and S.G. Churye, November 1976.
30. "Discrimination Between Stationary Gaussian Processes, Large Sample  
Results," Will Gersch, January 1977.
31. "Principal Components in the Nonnormal Case: The Test for Sphericity,"  
Christine M. Waternaux, October 1977.
32. "Nonnegative Definiteness of the Estimated Dispersion Matrix in a  
Multivariate Linear Model," F. Pukelsheim and George P.H. Styan, May 1978.

TECHNICAL REPORTS (continued)

33. "Canonical Correlations with Respect to a Complex Structure," Steen A. Andersson, July 1978.
34. "An Extremal Problem for Positive Definite Matrices," T.W. Anderson and I. Olkin, July 1978.
35. "Maximum likelihood Estimation for Vector Autoregressive Moving Average Models," T. W. Anderson, July 1978.
36. "Maximum likelihood Estimation of the Covariances of the Vector Moving Average Models in the Time and Frequency Domains," F. Ahrabi, August 1978.
37. "Efficient Estimation of a Model with an Autoregressive Signal with White Noise," Y. Hosoya, March 1979.
38. "Maximum Likelihood Estimation of the Parameters of a Multivariate Normal Distribution," T.W. Anderson and I. Olkin, July 1979.
39. "Maximum Likelihood Estimation of the Autoregressive Coefficients and Moving Average Covariances of Vector Autoregressive Moving Average Models," Fereydoon Ahrabi, August 1979.
40. "Smoothness Priors and the Distributed Lag Estimator," Hirotugu Akaike, August, 1979.
41. "Approximating Conditional Moments of the Multivariate Normal Distribution," Joseph G. Deken, December 1979.
42. "Methods and Applications of Time Series Analysis - Part I: Regression, Trends, Smoothing, and Differencing," T.W. Anderson and N.D. Singpurwalla, July 1980.
43. "Cochran's Theorem, Rank Additivity, and Tripotent Matrices." T.W. Anderson and George P.H. Styan, August, 1980.
44. "On Generalizations of Cochran's Theorem and Projection Matrices," Akimichi Takemura, August, 1980.
45. "Existence of Maximum Likelihood Estimators in Autoregressive and Moving Average Models," T.W. Anderson and Raúl P. Mentz, Oct. 1980.
46. "Generalized Correlations in the Singular Case," Ashis Sen Gupta, November 1980.
47. "Updating a Discriminant Function on the Basis of Unclassified Data," G.J. McLachlan and S. Ganesalingam, November 1980.
48. "A New Proof of Admissibility of Tests in the Multivariate Analysis of Variance." T.W. Anderson and Akimichi Takemura, January, 1981.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE                                                                                                                                                                                                                                                                                                                                                                                  |                                     | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM                                   |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|-------------------------------------------------------------------------------|
| 1. REPORT NUMBER<br>48                                                                                                                                                                                                                                                                                                                                                                                     | 2. GOVT ACCESSION NO.<br>AD-A097110 | 3. RECIPIENT'S CATALOG NUMBER                                                 |
| 4. TITLE (and Subtitle)<br>A NEW PROOF OF ADMISSIBILITY OF TESTS IN THE<br>MULTIVARIATE ANALYSIS OF VARIANCE                                                                                                                                                                                                                                                                                               |                                     | 5. TYPE OF REPORT & PERIOD COVERED<br>Technical Report                        |
|                                                                                                                                                                                                                                                                                                                                                                                                            |                                     | 6. PERFORMING ORG. REPORT NUMBER                                              |
| 7. AUTHOR(s)<br>T. W. ANDERSON and AKIMICHI TAKEMURA                                                                                                                                                                                                                                                                                                                                                       |                                     | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-75-C-0442                            |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Department of Statistics<br>Stanford University<br>Stanford, California                                                                                                                                                                                                                                                                                     |                                     | 10. PROGRAM ELEMENT PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>(NR-042-034) |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research<br>Statistics and Probability Program Code 436<br>Arlington, Virginia 22217                                                                                                                                                                                                                                                            |                                     | 12. REPORT DATE<br>JANUARY 1981                                               |
|                                                                                                                                                                                                                                                                                                                                                                                                            |                                     | 13. NUMBER OF PAGES<br>29                                                     |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)                                                                                                                                                                                                                                                                                                                                |                                     | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED                          |
|                                                                                                                                                                                                                                                                                                                                                                                                            |                                     | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE                                 |
| 16. DISTRIBUTION STATEMENT (of this Report)<br><br>APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.                                                                                                                                                                                                                                                                                                    |                                     |                                                                               |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)                                                                                                                                                                                                                                                                                                                 |                                     |                                                                               |
| 18. SUPPLEMENTARY NOTES                                                                                                                                                                                                                                                                                                                                                                                    |                                     |                                                                               |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br><br>MANOVA, admissibility, majorization, Schur-convexity, extreme points,<br>exponential family.                                                                                                                                                                                                                     |                                     |                                                                               |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br><br>A new proof of admissibility of tests in MANOVA is given using Stein's<br>theorem (1956). The convexity condition of Stein's theorem is proved directly<br>by means of majorization rather than by the supporting hyperplane approach.<br>This makes the geometrical meaning of the admissibility result clearer. |                                     |                                                                               |

DD FORM 1473  
1 JAN 73EDITION OF 1 NOV 68 IS OBSOLETE  
S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

