A MONOTONICITY PROPERTY OF THE POWER FUNCTIONS OF SOME INVARIANT TESTS FOR MANOVA[1]

by

Morris L. Eaton and Michael D. Perlman

Technical Report No. 193

December 1972

University of Minnesota
Minneapolis, Minnesota

[1]Research supported in part by National Science Foundation Grant GP-34482.

Abstract

A Monotonicity Property of the Power Functions of Some Invariant Tests for MANOVA

The main result of the current research describes a monotonicity property of certain invariant tests for the multivariate analysis of variance problem. Suppose $X: r \times p$ has a normal distribution, $EX = \Theta$ and the rows of X are independent, each with covariance matrix $\Sigma: p \times p$. If K is the acceptance region of an invariant test, let $\rho_K(\delta)$ denote the power function of K, where $\delta = (\delta_1, \ldots, \delta_t)$, $t \equiv \min(r,p)$ and $\delta_1^2, \ldots, \delta_t^2$ are the t largest characteristic roots of $\Theta\Sigma^{-1}\Theta'$. A main result is

<u>Theorem.</u>

If K is a convex set (in (X, S)), then $\rho_K(\delta)$ is a Schur-convex function of δ.

Standard tests to which the above theorem can be applied include the Roy maximum root test and the Lawley-Hotelling trace test.

American Mathematical Society 1970 subject classification: Primary 62H15.

Key words and phrases: monotonicity, power function, invariant tests, MANOVA, convex set, G-orbit, Schur-convex.

1. Introduction.

We begin by considering a canonical form of the multivariate analysis of variance (MANOVA) testing problem suitable for studying the power functions of invariant tests (see Anderson (1958), Chap. 8). Suppose $X: r \times p$ is a random matrix whose rows are independent normally distributed with common covariance matrix $\Sigma: p \times p$, and let $EX = \Theta$. Let $S: p \times p$ be independent of X and have the Wishart distribution--$S \sim W(\Sigma, p, n)$. It is assumed throughout that Σ is positive definite and $n \geq p$. The MANOVA problem in the form given here is to test $H_0: \Theta = 0$ against the alternative $H_1: \Theta \neq 0$.

The above testing problem is invariant under all transformations of the form

$$(1.1) \qquad (X, S) \to (\Gamma X A', ASA')$$

where $\Gamma: r \times r$ is an orthogonal matrix and $A: p \times p$ is a non-singular real matrix. Let $t = \min\{r, p\}$. A maximal invariant statistic is $(c_1, \ldots, c_t) \equiv (c_1(XS^{-1}X'), \ldots, c_t(XS^{-1}X'))$ where $c_1 \geq c_2 \geq \cdots \geq c_t \geq 0$ are the ordered t-largest characteristic roots of $XS^{-1}X'$. A maximal invariant parameter is $(\gamma_1, \ldots, \gamma_t)$ where $\gamma_1 \geq \cdots \geq \gamma_t \geq 0$ are the t-largest characteristic roots of $\Theta\Sigma^{-1}\Theta'$. Let $\delta_i = \sqrt{\gamma_i}$, $i = 1, \ldots, t$, and $\delta = (\delta_1, \ldots, \delta_t)'$.

Denote by $\mathcal{S}_p^+$ the space of $p \times p$ positive definite matrices and let $K \subseteq R^{rp} \times \mathcal{S}_p^+$. Define $\pi_K(\Theta, \Sigma)$ by

$$(1.2) \qquad \pi_K(\Theta, \Sigma) = P_{\Theta,\Sigma}\{(X, S) \notin K\}$$

so π_K is the power function of the test with acceptance region K. If K is invariant under the transformations (1.1), then $\pi_K(\Theta, \Sigma) = \pi_K(\Gamma\Theta A', A\Sigma A')$ so

(1.3) $\pi_K(\Theta, \Sigma) = \pi_K(\Delta(\delta), I) \equiv \rho_K(\delta)$

where $\Delta(\delta)$: $r \times p$ satisfies $\Delta_{ii}(\delta) = \delta_i$ for $i = 1,\ldots, t$ and the remaining elements of $\Delta(\delta)$ are zero.

Definition 1.1.

Let $\mathcal{K}_1$ be the class of regions $K \subseteq R^{rp} \times \mathcal{S}_p^+$ such that

(a) K is invariant under all the transformations (1.1),

(b) K is convex in each row of X when S and the remaining rows of X are fixed.

Das Gupta, Anderson, and Mudholkar (1964) established the following theorem.

Theorem 1.1.

If $K \in \mathcal{K}_1$, then $\rho_K(\delta)$ is increasing in each δ_i, $i = 1,\ldots, t$.

The following well known acceptance regions are in $\mathcal{K}_1$.

(i) Roy's Maximum root test:

$$K_1 = \{(X, S) | c_1(XS^{-1}X') \leq k\}, \ k > 0$$

(ii) Lawley-Hotelling trace test:

$$K_2 = \{(X, S) | \operatorname{tr} XS^{-1}X' = \sum_1^t c_i \leq k\}, \ k > 0$$

(iii) Likelihood Ratio Test:

$$K_3 = \{(X, S) | \prod_{i=1}^t (1 + c_i) \leq k\}, \ k > 0$$

(iv) Pillai's trace test:

$$K_4 = \{(X, S) | \operatorname{tr} X(X'X + S)^{-1}X' = \sum_1^t \frac{c_i}{1+c_i} \leq k\}, \ 0 < k \leq 1.$$

Although Theorem 1.1 shows that $\rho_{K_j}(\delta)$ is increasing in each δ_i, $j = 1,\ldots, 4$, this result does not help one choose among the four tests if high power at certain alternatives is desired. Numerical studies have been provided by Pillai and Jayachandran (1967, 1968) when $t = 2$ and by Fujikoshi (1970) when $t = 3$ which allow the study of the behavior of these power functions along certain contours in the alternative space. It is the purpose of this paper to provide an initial theoretical result concerning the behavior of $\rho_{K_j}(\delta)$ for $j = 1, 2$ along contours linear in δ.

2. The Main Theorem.

The discussion in this section is concerned with the following subclass of $\mathcal{K}_1$.

Definition 2.1.

Let $\mathcal{K}_2$ be the class of regions $K \subseteq R^{rp} \times S_p^+$ such that

(a) K is invariant under all the transformations (1.1),

(b') K is a convex set.

Clearly, $\mathcal{K}_2 \subseteq \mathcal{K}_1$. It will be shown (Theorem 4.1) that $K_1, K_2 \in \mathcal{K}_2$, but it is easy to verify that $K_3, K_4 \notin \mathcal{K}_2$.

Let $\lambda = (\lambda_1, \ldots, \lambda_t)'$ and define $\Delta(\lambda)$ to be the $r \times p$ real matrix with $\Delta_{ii}(\lambda) = \lambda_i$, $i = 1, \ldots, t$, and the other elements of $\Delta(\lambda)$ are zero. Then extend ρ_K by

$$\rho_K(\lambda) = \pi_K(\Delta(\lambda), I).$$

Let G_0 be the group of permutations and sign changes of coordinates acting on R^t. It is easy to verify that $\rho_K(\lambda) = \rho_K(g\lambda)$ for $g \in G_0$. For $\lambda_0 \in R^t$, the G_0-orbit of λ_0 is defined by

$$\mathfrak{G}(\lambda_0) = \{g\lambda_0 | g \in G_0\}.$$

Also, let $C(\lambda_0)$ be the convex hull of $\mathfrak{G}(\lambda_0)$. The following result is our main theorem.

Theorem 2.1.

If $K \in \mathcal{K}_2$ and $\lambda_0 \in R^t$, then

$$\rho_K(\lambda) \leq \rho_K(\lambda_0)$$

for all $\lambda \in C(\lambda_0)$.

The proof of this theorem is given in the next section. The following diagram illustrates and compares Theorems 1.1 and 2.1 when $t = 2$.

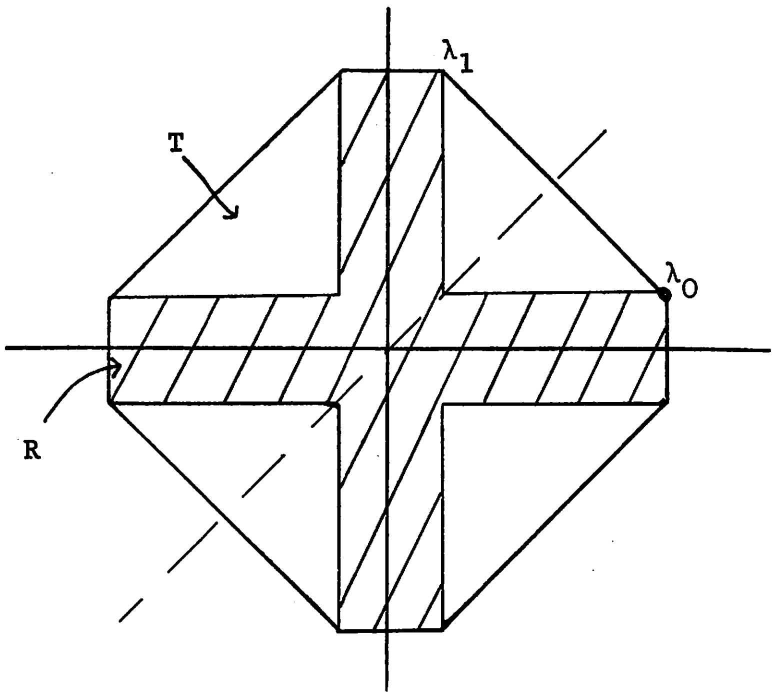

Figure 1.

Theorem 1.1 shows that if $K \in \mathcal{K}_1$, then $\rho_K(\lambda) \leq \rho_K(\lambda_0)$ for λ in the shaded region R. Our Theorem 2.1 shows that if $K \in \mathcal{K}_2$, then $\rho_K(\lambda) \leq \rho_K(\lambda_0)$ for all $\lambda \in T$. Note that the vertices of the octagon T are the points in the G_0-orbits of λ_0, and $T = C(\lambda_0)$.

Let GP_0 be the permutation subgroup of G_0, $\tilde{\mathfrak{S}}(\lambda_0) = \{g\lambda_0 | g \in GP_0\}$ and $\tilde{C}(\lambda_0)$ be the convex hull of $\tilde{\mathfrak{S}}(\lambda_0)$. It follows from Theorem 2.1 that

$$\rho_K(\lambda) \leq \rho_K(\lambda_0), \quad \lambda \in \tilde{C}(\lambda_0).$$

(In Fig. 1, $\tilde{\mathcal{O}}(\lambda_0) = \{\lambda_0, \lambda_1\}$ and $\tilde{C}(\lambda_0)$ is the line segment $[\lambda_0, \lambda_1]$.) But, $\lambda \in \tilde{C}(\lambda_0)$ if and only if $\lambda = Q\lambda_0$ for some doubly stochastic matrix Q, i.e., λ_0 <u>majorizes</u> λ (Berge, 1963, Chap. 8) so

(2.1) $$\rho_K(Q\lambda_0) \leq \rho_K(\lambda_0)$$

for all doubly stochastic matrices Q. But, (2.1) is the definition of Schur-convexity (Berge (1963), p. 219). Thus, we have proved

<u>Corollary 2.1.</u>

If $K \in \mathcal{K}_2$, then $\rho_K(\lambda)$ is a Schur-convex function of λ.

3. <u>Proof of Theorem 2.1.</u>

The proof of Theorem 2.1 is based on the following result, due to Mudholkar (1966). Let G be a group of Lebesgue measure preserving linear transformations on R^m. If $x \in R^m$, let $\mathcal{O}(x) = \{gx | g \in G\}$ be the orbit of x and let $C(x)$ be the convex hull of $\mathcal{O}(x)$.

<u>Theorem 3.1.</u>

Let f be a probability density on R^m such that $f(x) = f(gx)$ for all $x \in R^m$ and $g \in G$, and assume f is unimodal, i.e., $\{x | f(x) \geq c\}$ is convex for all $c > 0$. Further, let E be a convex set in R^m such that $E = gE$ for $g \in G$. Then

$$\int_E f(x+y)dx \geq \int_E f(x+y_0)dx$$

for all $y \in C(y_0)$.

To apply Theorem 3.1 to the MANOVA problem, let G be the group of transformations (Γ_1, Γ_2) acting on points $(X, S) \in R^{rp} \times R^{\frac{1}{2}(p(p+1))} = R^m$ by

(3.1) $$(\Gamma_1, \Gamma_2)(X, S) = (\Gamma_1 X \Gamma_2', \Gamma_2 S \Gamma_2')$$

where Γ_1: $r \times r$, Γ_2: $p \times p$ are orthogonal. Points in $R^{\frac{1}{2}(p(p+1))}$ are represented as real symmetric $p \times p$ matrices S. Theorem 2.1 is a consequence of the following lemma.

Lemma.

Suppose $K \in \mathcal{K}_2$. Then

$$\pi_K(\Lambda, I) \leq \pi_K(\Lambda_0, I)$$

for all Λ such that $(\Lambda, 0) \in C(\Lambda_0, 0)$, where $(\Lambda, 0) \in R^{rp} \times R^{\frac{1}{2}(p(p+1))}$.

Proof:

We will apply Theorem 3.1 with

$$f(X, S) = c|S|^{\frac{1}{2}(n-p-1)} \exp[-\tfrac{1}{2} \operatorname{tr}(X'X + S)]I(S)$$

where $I(S)$ is the indicator function of the set of positive definite matrices in $R^{\frac{1}{2}(p(p+1))}$. Clearly f is invariant under the transformations (3.1) and K is convex and invariant by assumption. The unimodality of f is verified by noting that $\log f(X, S)$ is a concave function of (X, S). Thus, by Theorem 3.1,

$$\int_K f((X, S) - (\Lambda, 0))dXdS \geq \int_K f((X, S) - (\Lambda_0, 0)dXdS$$

for $(\Lambda, 0) \in C(\Lambda_0, 0)$. But,

$$1 - \pi_K(\Lambda, I) = \int_K f((X, S) - (\Lambda, 0))dXdS,$$

and the result follows.

Proof of Theorem 2.1.

We give the proof for $r \leq p$ so $t = r$, as the case $p > r$ is similar. It is easy to show that the group G_0 is the set of all $r \times r$ matrices of the form $D_\varepsilon P$ where $D_\varepsilon : r \times r$ is a diagonal matrix with diagonal entries ± 1 and $P: r \times r$ is a permutation matrix. The relations

$$\Delta(D_\varepsilon \lambda) = D_\varepsilon \Delta(\lambda) \tag{3.2}$$

$$\Delta(P\lambda) = P\Delta(\lambda)\begin{pmatrix} P & 0 \\ 0 & I_{p-r} \end{pmatrix} \tag{3.3}$$

are easily verified.

Suppose $\lambda \varepsilon C(\lambda_0)$--i.e.,$\lambda$ is in the convex hull of the G_0-orbit of λ_0. Thus λ has the form

$$\lambda = \sum_i \alpha_i D_{\varepsilon,i} P_i \lambda_0$$

where $0 \leq \alpha_i \leq 1$ and $\Sigma\alpha_i = 1$. Using (3.2), (3.3), and the linearity of $\Delta(\cdot)$,

$$\Delta(\lambda) = \Delta(\Sigma\alpha_i D_{\varepsilon,i} P_i \lambda_0) = \Sigma\alpha_i D_{\varepsilon,i} P_i \Delta(\lambda_0)\begin{pmatrix} P_i & 0 \\ 0 & I_{p-r} \end{pmatrix}.$$

Since $D_{\varepsilon,i}P_i$: $r \times r$ and $\begin{pmatrix} P_i & 0 \\ 0 & I_{p-r} \end{pmatrix}$: $p \times p$ are orthogonal, this shows that $(\Delta(\lambda), 0)$ is in the convex hull of the G orbit of $(\Delta(\lambda_0), 0)$. Applying the lemma, $\rho_K(\lambda) = \pi_K(\Delta(\lambda), I) \leq \pi_K(\Delta(\lambda_0), I) = \rho_K(\lambda_0)$. This completes the proof.

4. Applications of Theorem 2.1 to Specific Tests.

In this section, we give a sufficient condition for an invariant acceptance region to be in $\mathcal{K}_2$.

Definition 4.1.

Let $R^t_+ = \{w \varepsilon R^t | w_i \geq 0, i = 1,\ldots, t\}$ and let $\mathfrak{W}$ be the class of regions $W \subseteq R^t_+$ satisfying

(i) W is closed and convex

(ii) W is invariant under permutation of coordinates

(iii) W is monotone: if $w \varepsilon W$ and if $0 \leq v_i \leq w_i$, $i = 1,\ldots, t$, then $v \varepsilon W$.

Let $K_W = \{(X, S) | (c_1,\ldots, c_t) \varepsilon W\}$ where $c_1,\ldots, c_t$ are the t-largest ordered characteristic roots of $XS^{-1}X'$.

Theorem 4.1.

If $W \varepsilon \mathfrak{W}$, then $K_W \varepsilon \mathcal{K}_2$.

Proof:

K_W is clearly invariant under the transformations (1.1). We outline the proof of the convexity of K_W (see Schwartz (1967) for details). Since $W \in \mathfrak{W}$, W can be represented as an intersection of sets of the form

$$\{(X, S) | \sum_{i=1}^{t} b_i c_i(XS^{-1}X') \leq 1\} \equiv K_b$$

where $b_1 \geq b_2 \geq \ldots \geq b_t \geq 0$. Let $D_w = \text{diag}\{w_1, \ldots, w_t\}$. Since $\Sigma b_i c_i = \text{tr } D_b D_c = \sup_{\Psi} \text{tr } D_b \Psi X S^{-1} X' \Psi'$ where $\Psi: t \times r$ ranges over all row orthogonal matrices, it is clear that $(X, S) \in K_b$ iff

$$\text{tr } D_b \Psi X S^{-1} X' \Psi' \leq 1 \quad \text{for all } \Psi.$$

From the Cauchy-Schwartz inequality we have for $F: t \times p$ and $H: t \times p$

$$\sup_{F \neq 0} \frac{(\text{tr } FH')^2}{\text{tr } FSF'} = \text{tr } H'HS^{-1} .$$

Setting $H = D_b^{\frac{1}{2}} \Psi X$, we see that $(X, S) \in K_b$ iff

$$(\text{tr } FX'\Psi' D_b^{\frac{1}{2}})^2 \leq \text{tr } FSF' \quad \text{for all } \Psi, F \neq 0.$$

Thus

$$K_b = \bigcap_{\Psi} \bigcap_{F} \{(X, S) | (\text{tr } FX'\Psi' D_b^{\frac{1}{2}})^2 - \text{tr } FSF' \leq 0\}.$$

But, $(\text{tr } FX'\Psi' D_b^{\frac{1}{2}})^2 - \text{tr } FSF'$ is a convex function of (X, S) so K_b is the intersection of convex sets. This completes the proof.

The following corollary is now immediate.

Corollary 4.1.

The regions K_1 and K_2 are in $\mathcal{K}_2$.

Let $\delta^{(1)} = (1, 0, \ldots, 0)$, $\delta^{(2)} = (\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0), \ldots, \delta^{(t)} = (\frac{1}{t}, \ldots, \frac{1}{t})$. Since $\delta^{(j+1)}$ is in the convex hull of the GP_0-orbit of $\delta^{(j)}$, we have from Corollary 2.1,

$$\rho_K(\delta^{(j+1)}) \leq \rho_K(\delta^{(j)}),\ K \in \mathcal{K}_2.$$

When $\delta = \delta^{(j)}$, the rank of Θ is j. Thus we conclude that for $\sum_{i=1}^{t} \delta_i =$ constant, the power of tests in $\mathcal{K}_2$ increases as the rank of Θ decreases from t to 1. This result confirms speculation based on the numerical results of Pillai and Jayachandran (1967, 1968) and Fujikoshi (1970). (These authors studied the power functions for $\Sigma\delta_i^2$ = constant.)

These numerical studies suggest the conjecture that the power functions of the tests K_3 and K_4 decrease as the rank of Θ decreases from t to 1. However the methods of the present paper cannot be applied as neither the acceptance nor rejection regions of the tests based on K_3 and K_4 are convex sets in (X,S).

One possible application of Corollary 2.1 is to derive upper and lower bounds for the power functions $\rho_K(\delta)$ when $K \in \mathcal{K}_2$. Since $\delta^* \equiv (\Sigma\delta_i, 0, \ldots, 0)$ majorizes δ and δ majorizes $\delta_* \equiv (t^{-1}\Sigma\delta_i, \ldots, t^{-1}\Sigma\delta_i)$,

$$\rho_K(\delta_*) \leq \rho_K(\delta) \leq \rho_K(\delta^*).$$

This provides an upper and lower envelope for $\rho_K(\delta)$ in terms of $\Sigma\delta_i$.

References

[1] Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley, New York.

[2] Berge, C. (1963). Topological Spaces. Macmillan, New York.

[3] Das Gupta, S., Anderson, T. W., and Mudholkar, G. S. (1964). Monotonicity of the power functions of some tests of the multivariate linear hypothesis. Ann. Math. Statist. 35 200-205.

[4] Fujikoshi, Y. (1970). Asymptotic expansions of the distributions of test statistics in multivariate analysis. J. Sci. Hiroshima Univ. (Series A-I) 34 73-144.

[5] 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.

[6] Pillai, K. L. S. and Jayachandran, K. (1967). Power comparisons of tests of two multivariate hypotheses based on four criteria. Biometrika 54 195-210.

[7] Pillai, K. L. S. and Jayachandran, K. (1968). Power comparisons of tests of equality of covariance matrices based on four criteria. Biometrika 55 335-342.

[8] Schwartz, R. E. (1967). Admissible tests in multivariate analysis of variance. Ann. Math. Statist. 38 698-710.