PROPERTIES OF TESTS CONCERNING COVARIANCE MATRICES OF NORMAL DISTRIBUTIONS

by

Somesh Das Gupta* and N. C. Giri

University of Minnesota   Universite de Montreal

Technical Report No. 150

February 1971

University of Minnesota
Minneapolis, Minnesota

*This research was supported by National Science Foundation Grant GP-21074.

1. Introduction and summary.

The problem of testing $\Sigma = \Sigma_0$ in $N_p(\mu, \Sigma)$ is considered in Section 2. A class of tests based on $|S|^{r/2}\mathrm{etr}(-S/2)$ is studied, where $S = \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(X_\alpha - \bar{X})'$, $\bar{X} = \sum_{\alpha=1}^{N} X_\alpha / N$ and X_α's constitute a random sample from $N_p(\mu, \Sigma)$. The first theorem deals with the questions of unbiasedness and monotonicity of the power functions. The second theorem deals with the question of admissibility and the third theorem presents a locally best invariant test for some special alternatives.

In the next section the problem of testing the equality of covariance matrices of two p-variate normal distributions is dealt with. A class of critical regions based on $|S_1|^a|S|^{b-a}/|S_1 + S_2|^b$ is considered where

$$S_1 = \sum_{\alpha=1}^{N_1} (X_\alpha - \bar{X})(X_\alpha - \bar{X})', \quad \bar{X} = \sum_{\alpha=1}^{N_1} X_\alpha / N_1$$

$$S_2 = \sum_{\alpha=1}^{N_2} (Y_\alpha - \bar{Y})(Y_\alpha - \bar{Y})', \quad \bar{Y} = \sum_{\alpha=1}^{N_2} Y_\alpha / N_2$$

and X_α's and Y_α's constitute random samples from $N_p(\mu_1, \Sigma_1)$ and $N_p(\mu_2, \Sigma_2)$, respectively. These tests have been studied from the viewpoint of unbiasedness.

The problem of testing the equality or proportionality of two or more covariance matrices is discussed in the last section. Some of the results in this paper are generalizations of those in [3] and [9].

2. Tests of $\Sigma = \Sigma_0$ in $N_p(\mu, \Sigma)$.

For testing $\Sigma = I_p$ (the hypothesis $\Sigma = \Sigma_0$, a specified matrix, can be reduced to $\Sigma = I_p$) in $N_p(\mu, \Sigma)$ several heuristic tests have been proposed in the literature and some properties of these test have been studied. In this section we consider a class of tests from the point of view of unbiasedness and monotonicity of the power function and later study the question of admissibility and local optimal property.

Let $S = \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(X_\alpha - \bar{X})' = [s_{ij}]$, $\bar{X} = \sum_{\alpha=1}^{N} X_\alpha / N$,

where $X_1, \ldots, X_N$ constitute a random sample from $N_p(\mu, \Sigma)(N > p)$. The sufficiency and the invariance (with respect to the direct product of the translation group and the group of $p \times p$ orthogonal matrices)lead to the class of tests based on the characteristic roots $c_1, \ldots, c_p$ of S. The first general result, quoted below, was obtained by Anderson and Das Gupta [1].

Theorem.

Let w be a region in the space of $c_1, \ldots, c_p$ such that $(c_1, \ldots, c_p) \in w \Rightarrow (c_1^-, \ldots, c_p^-) \in w$ where $c_i^- \leq c_i$. Then $P(w|\Sigma)$ is a function of the characteristic roots $\gamma_1, \ldots, \gamma_p$ of Σ and decreases as each γ_i increases.

The above result can only be applied to one-sided problems and, even in that case nothing generally is known about $P(w|\Sigma)$ when some γ_i's increase and some others decrease. For the two-sided alternatives (i.e., $\Sigma \neq I_p$) the likelihood-ratio test (LRT) was shown to be biased [3] and the modified likelihood-ratio test (MLRT) was not only shown to be unbiased [9] but also to enjoy the monotonicity property [3].

These results lead to the consideration of a general class of critical regions given by

$$w_r:\ |S|^{r/2}\,\mathrm{etr}(-S/2) \leq \lambda \tag{2.1}$$

for $r \geq 0$; when $r \leq 0$ w_r^c, the complement of w_r, satisfies the condition in Anderson-Das Gupta's theorem.

Using the methods in [3, 9] we get the following theorem. Let $\Gamma = \mathrm{diag}\ (\gamma_1, \ldots, \gamma_p)$. Without loss of generality, we shall assume $\Sigma = \Gamma$. $n = N - 1$.

<u>Theorem 2.1.</u>

(a) $P(w_r|\Gamma)$ <u>increases monotonically as each</u> γ_i <u>deviates from</u> 1 <u>either in the positive direction or in the negative direction.</u>

(b) <u>The test of</u> $H_{01}: \Sigma = I_p$ <u>with the critical region</u> w_r, <u>given by (2.1), is biased for</u> $1 < \gamma_i < r/n$ $(i = 1, \ldots, n)$ <u>if</u> $n < r$ <u>and for</u> $r/n < \gamma_i < 1$ $(i = 1, \ldots, n)$ <u>if</u> $r < n$.

(c) <u>The test of</u> H_{01} <u>with the critical region</u> w_r <u>is unbiased for the alternatives</u> $|\Sigma| \leq 1$ <u>if</u> $n \leq r$ <u>and for</u> $|\Sigma| \geq 1$ <u>if</u> $r \leq n$.

(d) $P(w_r|\Gamma) \to 1$ <u>if any</u> γ_i <u>tends to</u> 0 <u>or</u> ∞, <u>when</u> $r > 0$.

<u>Proof</u>:

(a) It is easy to show that if Y is distributed as $\gamma\chi_n^2 (\gamma > 0)$ then $P(Y^{r/2} \exp(-Y/2) \leq \lambda)$ monotonically decreases for $\gamma \in (0, r/n)$ and increases for $\gamma \in (r/n, \infty)$. Note that

$$|S|^{r/2}\mathrm{etr}(-S/2) = \{|S|/\prod_{i=1}^{p} s_{ii}\}^{r/2}\{\prod_{i=1}^{p} s_{ii}^{r/2}\exp(-s_{ii}/2)\}. \tag{2.2}$$

Since the distribution of the first factor in the right hand side of (2.2) is free of γ_i's, it is independent of the second factor. Moreover, S_{ii}'s are mutually independent. From the above facts (a) follows easily.

(b) Follows from (a).

(c) Using the method of Suguira and Nagao [9] it can be shown that

$$P(w_r^c|\Gamma = I_p) - P(w_r^c|\Gamma) \tag{2.3}$$

$$\geq K(n, p)\exp(-\lambda/2)\{1 - |\Gamma|^{(r-n)/2}\} \int_{w_r^c} |S|^{(n-1-p-r)/2}dS$$

where $K(n, p)$ is a constant depending on n and p. The result (b) follows after noting that the integral in (2.3) is finite.

(d) Using Chebyshev's inequality,

$$P(w_r^c|\Gamma) = P[|S|^{r/2}\mathrm{etr}(-S/2) \leq \lambda|\Gamma] \leq (1/\lambda)E[|S|^{r/2}\mathrm{etr}(-S/2)|\Gamma]$$

$$= K(n, r, p, \lambda) \prod_{i=1}^{p} \{\gamma_i^{r/2}/(1+\gamma_i)^{(n+r)/2}\}$$

which tends to 0 if γ_i tends to 0 or ∞.

Nest, we use the technique of Kiefer and Schwartz [7] to examine the admissibility of a class of critical regions for testing $\Sigma \neq I_p$.

<u>Theorem 2.2.</u>

<u>For testing</u> $\Sigma = I_p$ <u>against</u> $\Sigma \neq I_p$ <u>the following critical regions are unique (a.e.) Bayes and hence admissible when</u> $n > p$.

(i) $|S|^{r/2}\mathrm{etr}(-S/2) \leq \lambda,\ 1 < r < \infty$

(ii) $|S|^{r/2}\mathrm{etr}(-S/2) \geq \lambda,\ -\infty < r < 0.$

Proof:

Consider the following prior distribution on the parameter space under the alternative (μ and $\bar{X}$ can be eliminated as in [7]):

$$\Sigma^{-1} = cI_p + \eta\eta', \ c > 0$$

where the density (with respect to the Lebesgue measure) of η: $p \times 1$ is proportional to

$$|cI_p + \eta\eta'|^{-n/2}.$$

The above function is integrable if $n > p$ [7]. The Bayes critical region is given by

$$|S|^{1/2} \mathrm{etr}\{- (1-c)S/2\} \leq \lambda.$$

(i) is obtained for $r = 1/(1-c)$, $0 < c < 1$ and (iii) for $r = 1/(1-c)$, $1 < c$.

Remark.

The critical region given by $\mathrm{tr}(S) \leq \lambda$ (or, $\geq \lambda$) can be seen to be Bayes against $\Sigma = cI_p$ when $1 < c$ (or, $c < 1$). For two-sided tests based on $\mathrm{tr}(S)$, see [7]. The other cases (i.e., for other values of r) cannot be studied by this method due to the special structure of Σ^{-1} and these cases are yet to be explored.

Next, we consider the problem of finding a locally best invariant test of $\Sigma = I_p$. It will be indicated in the following discussion that the nature of the alternative hypothesis plays a vital role in getting a locally best invariant test. Suppose we want to test $\theta = \theta^o$ against $\theta \in \Theta_1$ where θ is a real $r \times 1$ vector. Following Lehmann [8] a locally best level α test may be defined as a test φ_0 such that given any other level α test φ there exists $\Delta > 0$ for which

(2.5) $\quad 0 < d(\theta) < \Delta$ and $\theta \in \Theta_1 \Rightarrow E_\theta(\varphi) \leq E_\theta(\varphi_0)$

where d is a measure of the distance of θ from θ^o. Suppose now $\beta(\theta, \varphi) = E_\theta(\varphi)$ has a differential at $\theta = \theta^o$ for every test φ. Then given any two tests φ and φ_0, and given $\varepsilon > 0$, there exists $\delta > 0$ such that

(2.6)
$$|\{\beta(\theta, \varphi_0) - \beta(\theta, \varphi)\} - \{\beta(\theta^o, \varphi_0) - \beta(\theta^o, \psi)\} - \sum_{i=1}^{r}(\theta_i - \theta_i^o)\frac{\partial}{\partial\theta_i}\{\beta(\theta^o, \varphi_0) - \beta(\theta^o, \varphi)\}| < \varepsilon\,\|\theta - \theta^o\|$$

if $0 < \|\theta - \theta^o\| < \delta$, where $\|\cdot\|$ denotes the Euclidean distance and $\theta' = (\theta_1, \ldots, \theta_r)$. In the one-parameter case, a solution is obtained by finding a size α test φ_0 which maximizes the derivative of the power function at $\theta = \theta^o$ among all level α tests when the alternatives are $\theta > \theta^o$. This idea cannot be used in the multiparameter case unless the problem has some special structure.

In the present context, let $\beta(\Gamma, \varphi) = E_\Gamma\varphi$ for a test φ depending on c_i's. The joint distribution of $c_1 < \ldots < c_p$ is [6]

$$K(n,p)|\Gamma|^{-n/2}\prod_{i>j}(c_i - c_j)\prod_{i=1}^{p}c_i^{(n-p-1)/2}\int_G \operatorname{etr}(-\Gamma^{-1}gcg'/2)d\nu(g)$$

where G is the class of all $p \times p$ orthogonal matrices and ν is the left-invariant probability measure on G. It can be seen that

(2.7)
$$2\frac{\partial}{\partial\gamma_i}\beta(\Gamma, \varphi)]_{\Gamma=I_p} = E[(-n + \sum_{j=1}^{p}c_j/p)\varphi|\Gamma = I_p]$$

for any test φ. Using (2.5), (2.6), (2.7) and the Neyman-Pearson lemma, we get the following theorem:

Theorem 2.3.

The locally best invariant test of $\Sigma = I_p$ against $tr(\Sigma) > tr(I_p)$ is the size α test having the critical region $tr(S) > k$; for the alternatives $tr(\Sigma) < tr(I_p)$, it is the size α test having the critical region $tr(S) < k$.

Remark.

The alternatives $tr(\Sigma) > p$ (or, $< p$) may look artificial but only with this kind of alternatives the problem of finding a locally best invariant test yields a solution.

We have been unable to prove whether the tests in Theorem 2.3 are unbiased against the respective alternatives. All we can say is that the critical region $tr(S) > k$ is unbiased for $|\Sigma| \geq 1$ which is included in the region $tr(\Sigma) \geq p$.

The problem of finding locally best test is related to that of local minimax test [5].

3. Tests of the equality of two covariance matrices.

Let $S_1 = [s_{ij}^{(1)}]$ and $S_2 = [s_{ij}^{(2)}]$ be matrices as defined in the introduction; let $n_i = N_i - 1$. The first general result, quoted below, regarding the monotonicity of the power functions of tests of $\Sigma_1 = \Sigma_2$ was obtained by Anderson and Das Gupta [1].

Theorem.

Let w be a region in the space of the characteristic roots $c_1, \ldots, c_p$ of $S_1 S_2^{-1}$ such that $(c_1, \ldots, c_p) \in w \Rightarrow (c_1^-, \ldots, c_p^-) \in w$, $c_i^- \leq c_i$. Then $P(w | \Sigma_1, \Sigma_2)$ depends only on the characteristic roots $\gamma_1, \ldots, \gamma_p$ of $\Sigma_1 \Sigma_2^{-1}$ and decreases as each γ_i increases.

Again this theorem gives a partial answer to the nature of $P(w|\Sigma_1, \Sigma_2)$ when γ_i's vary and it is applicable only to one-sided problems.

Suguira and Nagao [9] proved that the MLRT is unbiased against $\Sigma_1 \neq \Sigma_2$ and Das Gupta [3] showed that the LRT is biased when $n_1 \neq n_2$. Here we consider a general class of critical regions given by

$$(3.1) \qquad w(a, b): \ |S_1|^a |S_2|^{b-a} / |S_1 + S_2|^b \leq \lambda.$$

For the LRT, $a = n_1 + 1$, $b = n_1 + n_2 + 2$ and for the MLRT, $a = n_1$, $b = n_1 + n_2$. If $a(a-b) > 0$, Anderson-Das Gupta's theorem can be used to study $P(w(a, b)|\Sigma_1, \Sigma_2)$ regarding unbiasedness or monotonicity. In the following, we assume $0 < a < b$; in this case, the tests given by (3.1) are known to be admissible [7].

Theorem 3.1.

(a) *The critical region* $w(a, n_1 + n_2)$ *is unbiased for* $H_{02}: \Sigma_1 = \Sigma_2$ *against the alternatives* $|\Sigma_1| \geq |\Sigma_2|$ *when* $a \leq n_1$ *and against the alternatives* $|\Sigma_1| \leq |\Sigma_2|$ *when* $n_1 \leq a$.

(b) *The critical region* $w(a, b)$ *is biased for* H_{02} *against the alternatives* $1 \leq \gamma_i \leq d$ $(i = 1, \ldots, p)$ *when* $1 < d$ *and against the alternatives* $d \leq \gamma_i \leq 1$ $(i = 1, \ldots, p)$ *when* $d < 1$, *where* $d = a(n_1 + n_2)/bn_1$.

Proof:

(a) Without loss of generality, assume $\Sigma_1 = \Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_p)$, $\Sigma_2 = I_p$. Define $U = S_1^{-1/2} S_2 S_1^{-1/2}$. The density of U is given by [9]

$$(3.2) \qquad K(p, n_1, n_2)|\Gamma|^{-n_1/2}|U|^{(n_2-p-1)/2}|\Gamma^{-1}+U|^{-(n_1+n_2)/2}.$$

The region $w(a, b)$ can be expressed as

$$(3.3) \qquad |U|^{b-a}/|I_p+U|^b \leq \lambda.$$

Following the proof of Suguira and Nagao [9], we get

$$(3.4) \qquad \begin{aligned} &P(w^c(a, n_1+n_2)|H_{20}) - P(w^c(a, n_1+n_2)|\Gamma) \\ &\geq K(p, n_1, n_2)\lambda^{1/2}\{1 - |\Gamma|^{(a-n_1)/2}\iint_{w^c(a,n_1+n_2)} |U|^{(a-n_1-p-1)/2} dU. \end{aligned}$$

The desired result now follows from (3.4) after noting that the integral in (3.4) is finite.

(b) Consider a family of regions given by

$$R(a, b): y^a(1+y)^{-b} \geq k \ ; \ y > 0.$$

The region $R(a, b)$ is either an interval or the complement of an interval. When $0 < a < b$, $R(a, b)$ is a finite interval not including zero (excluding the trivial extreme case). The following lemma can be proved easily by differentiation.

<u>Lemma</u>.

Let Y be a random variable such that $Y/\delta (\delta > 0)$ is distributed as the ratio of two independent $\chi^2_{n_1}$ and $\chi^2_{n_2}$ variates. Let

$$\beta(\delta) = P[Y \in R(a, b)]$$

where $0 < a < b$. Then $\beta(\delta) \geq \beta(1)$ if δ lies between 1 and d and the strict inequality holds if δ lies in the open interval with endpoints 1 and d, where $d = a(n_1+n_2)/bn_1$.

Define Z by

$$(3.5) \qquad |S_1|^a |S_2|^{b-a} / |S_1 + S_2|^b = \{(s_{11}^{(1)})^a (s_{11}^{(2)})^{b-a} (s_{11}^{(1)} + s_{11}^{(2)})^{-b}\} Z.$$

Suppose $\Gamma = \mathrm{diag}(\gamma_1, 1, 1, \ldots, 1)$. Then the distribution of Z is free from γ_1 and independent of the first factor in the right-hand side of (3.5). It follows from the above lemma that the power of the critical region $w(a, b)$ at γ_1 is less than its size if γ_1 lies between 1 and d strictly. The result (b) now follows.

The problem of finding a locally best invariant test of $\Sigma_1 = \Sigma_2$ is discussed in[4].

4. Proportional covariance matrices.

Consider the problem of testing $\Sigma_1 = \Sigma_2$ against $\lambda\Sigma_1 = \Sigma_2$ $(\lambda \neq 1)$ for two normal distributions $N_p(\mu_1, \Sigma_1)$ and $N_p(\mu_2, \Sigma_2)$. Using sufficiency and invariance we consider only the class of tests based on the roots $c_1, \ldots, c_p$ of $S_1 S_2^{-1}$ where S_1, S_2 are defined in the introduction. The density of the roots $c_1 < \ldots < c_p$ under $\lambda\Sigma_1 = \Sigma_2$ is [6]

$$K(n_1, n_2, p) \prod_{i>j} (c_i - c_j) \prod_{i=1}^{p} c_i^{(n_1-p-1)/2} \lambda^{n_1 p/2} |I_p + \lambda C|^{-(n_1+n_2)/2}$$

$$= K(n_1, n_2, p) \prod_{i>j} (c_i - c_j) \prod_{i=1}^{p} c_i^{(n_1-p-1)/2} \lambda^{n_1 p/2} |I_p + C|^{-(n_1+n_2)/2}$$

$$\cdot \, |I_p + (\lambda-1) C (I_p + C)^{-1}|^{-(n_1+n_2)/2}$$

where $C = \mathrm{diag}(c_1, \ldots, c_p)$. Following Giri [4] or by using the method indicated in Section 2, it can be seen that the locally best invariant test of $\Sigma_1 = \Sigma_2$ against $\lambda\Sigma_1 = \Sigma_2$ $(\lambda > 1)$ has the critical region

$$\mathrm{tr}[C(I_p + C)^{-1}] \leq k$$

i.e.,

$$\mathrm{tr}[S_1(S_1 + S_2)^{-1}] \leq k.$$

It can be seen easily that the power of this test increases with λ.

The LRT for this problem is quite complicated. Even when $p = 2$, $n_1 = n_2$, the LRT has the critical region

$$(1 + c_1 + c_2 + c_1c_2)/(\sqrt{c_1} + \sqrt{c_2})^2 \leq k.$$

It may be noted that when $p = 2$, $n_1 = n_2$, the LRT of $\lambda\Sigma_1 = \Sigma_2$ against $\lambda\Sigma_1 \neq \Sigma_2$ is based on $(c_1 + c_2)/\sqrt{c_1c_2}$.

A related problem is to test the hypothesis $\Sigma_1 = \dots = \Sigma_k$ against $\Sigma_1 = \ell_2\Sigma_2 = \dots = \ell_k\Sigma_k$ where Σ_i's are covariance matrices of k p-variate normal distributions. This problem occurs in the analysis of variance components. When ℓ_i's are greater than 1, Chakravarty [2] has suggested the critical region $\mathrm{tr}(\sum_{j=2}^{k} S_jS_i^{-1}) \leq \lambda$. It can be easily seen that the power of this test increases with each ℓ_i.

References.

[1] Anderson, T. W. and Das Gupta, S. (1964). Monotonicity property of the power functions of some tests of the equality of two covariance matrices. Ann. Math. Stat. 35 1059-1063.

[2] Chakravorty, S. (1968). On the analysis of variance test in the multivariate variance component model. Cal. Stat. Assoc. Bull. 17 57-78.

[3] Das Gupta, S. (1969). Properties of power functions of some tests concerning dispersion matrices. Ann. Math. Stat. 40 697-702.

[4] Giri, N. (1968). On tests of the equality of two covariance matrices. Ann. Math. Stat. 39 275-277.

[5] Giri, N. and Kiefer, J. (1964). Local and asymptotic minimax properties of multivariate tests. Ann. Math. Stat. 35 21-35.

[6] James, A. T. (1964). Distributions of matrix variates and latent roots derived from normal spaces. Ann. Math. Stat. 35 475-501.

[7] Kiefer, J. and Schwartz, R. (1965). Admissible Bayes character of T^2-, R^2- and other fully invariant tests for classical multivariate normal problems. Ann. Math. Stat. 36 747-770.

[8] Lehmann, E. L. (1954). Testing Statistical Hypotheses. Wiley, New York.

[9] Suguira, N. and Nagao, H. (1968). Unbiasedness of some two criteria for the equality of one or two covariance matrices. Ann. Math. Stat. 39 1686-1692.