Statistica Sinica 2(1992), 221-236

ON THE INVARIANCE STRUCTURE OF THE ONE-SIDED TESTING PROBLEM FOR A MULTIVARIATE NORMAL MEAN

Takeaki Kariya and Arthur Cohen

Hitotsubashi University and Rutgers University

Abstract: This paper studies the invariant structure of the one-sided testing problem $\mu = 0$ vs $\mu \geq 0$ in $N_p(\mu, \Sigma)$ and it is shown that the requirement of scale invariance, symmetry in coordinates and similarity (transitivity) inevitably leads to the Hotelling T^2 test. We regard this as a negative result in the following sense: Since we cannot recommend Hotelling T^2 test for this one-sided problem it is fruitless to seek an invariant-similar test for this problem.

In addition to the above result, a noncentral bivariate t-distribution is derived, a different invariance approach is proposed when similarity is not obtained though the usual invariance approach, and the non-Bayes property of the Hotelling T^2 test is shown.

Key words and phrases: One-sided testing problem, bivariate t distribution, scale invariance, permutation invariance, similar test, Hotelling T^2.

1. Introduction

In this paper, from an invariance viewpoint, the well known one-sided testing problem

$$H: \eta = 0 \quad \text{vs} \quad K: \eta \geq 0, \ \eta \neq 0 \tag{1.1}$$

is studied in the p-dimensional normal model

$$x_1, \ldots, x_N \quad \text{iid} \quad N_p(\eta, \Sigma), \quad \Sigma \in S(p) \tag{1.2}$$

where $\eta \geq 0$ means $\eta_i \geq 0$, $i = 1, 2, \ldots, p$ for $\eta = (\eta_i)$ and $S(p)$ denotes the set of $p \times p$ positive definite matrices. This problem has been analyzed by various authors. Among others, Perlman (1969) studied it with a focus on some properties of the likelihood ratio test (LRT) and Eaton (1970) and Marden (1982) described essentially complete classes. See Perlman (1969) for many additional references. The LRT is not similar as shown in Perlman (1969). Tang (1990) gives similar tests more powerful than the likelihood ratio test. However his tests are not scale invariant. One of his tests is permutation invariant. The Hotelling T^2 test is an optimal invariant test for the two-sided problem $H : \eta = 0$ vs $K' : \eta \neq 0$. In this paper, we question, via invariance, whether there exists a nontrivial permutation invariant and scale invariant similar test except for the Hotelling T^2 statistic. The answer will be negative.

In general, a similar test is not necessarily invariant (under some group). However, similarity is often implied by invariance, and, most important, similar tests in practice are often invariant. That is, in most practical problems, the groups leaving the problems invariant act transitively on the null parameter space, which implies invariant tests are similar. From the traditional invariance viewpoint, transitivity will be the only way to deal with similarity at present, since no structural relationship between similarity and invariance has been established. Hence, in this paper similarity is viewed as the transitivity of a group action on the null parameter space by invariance; and we call a test invariantly similar if the similarity is guaranteed by the transitivity of the group action. It is shown that the Hotelling T^2 test is the unique nontrivial, coordinatewise scale invariant, permutation invariant and invariantly similar test for problem (1.1).

We regard our finding as a negative result in the following sense: We cannot seriously recommend Hotelling's T^2 test for this one-sided problem since it is very counter intuitive. It would reject, for example, when all of the p-sample means are very negative or even moderately negative, and, to boot, if the sample covariance matrix is near diagonal, with small diagonal elements! This is upsetting to any practical statistician. Furthermore, the power of Hotelling's T^2 test must suffer, relative to other tests, because of such rejections. Our result, then, implies it is fruitless to seek a test which simultaneously has the desirable properties of invariance and similarity. Further work needs to be done to determine which of the desirable properties should be sacrificed. Based on Tang's (1990) finding it appears that the LRT should not be used either.

In Section 2 the problem for the case $p = 2$ is studied. There, the nonnull distribution of a maximal invariant is derived, through which a noncentral bivariate t distribution is obtained as a by-product. In Section 3, we obtain a maximal group G_3 leaving the problem (1.1) invariant and including the group $P(p)$ of permutations and the group $A(p)$ of scale transformations. In general it is shown that a G_3 invariant test is not similar. Therefore, there is no group which leaves the problem (1.1) invariant that includes $P(p)$ and $A(p)$ and provides a class of similar tests as the class of invariant tests. For such a case, an alternative invariance approach, which we call an H–K invariance approach, is proposed in Section 4. In that approach we first choose a (maximal) group K leaving the nonnull model invariant, then choose a minimal group H containing K and acting transitively on the null parameter space, and finally choose an appropriate

H invariant test. In Section 5, this H–K invariance approach is applied to the problem. First, the minimal group which acts transitively on Θ_H (null space) and includes the group $A(p)$ of scale changes is shown to be the group $T(p)$ of $p \times p$ lower triangular matrices with positive diagonal elements. Second, the minimal group which acts transitively on Θ_H and includes the maximal group G_3 is shown to be the general linear group $G\ell(p)$ consisting of $p \times p$ nonsingular matrices. This result implies that the Hotelling T^2 test is the unique test which is permutation invariant, scale invariant and invariantly similar. In Section 6, it is shown that Hotelling T^2 test is not a proper Bayes test.

2. Testing $H : \eta = 0$ vs $K : \eta \geq 0$

In this section, to gain insight into the invariance structure of the problem, we directly analyze problem (1.1) when $p = 2$. By sufficiency, problem (1.1) is reduced to the problem of testing $H : \mu = 0$ vs $K : \mu \geq 0$ with $\mu = \sqrt{N}\eta$ in the model

$$\begin{cases} y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \sqrt{N}\bar{x}_1 \\ \sqrt{N}\bar{x}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma \right) \\ S = (s_{ij}) \sim W_2(\Sigma, n) \quad \text{with} \quad n = N - 1 \end{cases} \tag{2.1}$$

where y and S are independent and $W_2(\Sigma, n)$ represents a Wishart distribution. The problem is clearly left invariant under the scale group

$$G = \{A | A = \text{diag}\{a_1, a_2\}, \ a_i > 0 \ (i = 1, 2)\}, \tag{2.2}$$

under which a maximal invariant is

$$T = (t_1, t_2, r) \quad \text{with} \quad t_i = y_i / s_{ii}^{1/2} \quad \text{and} \quad r = s_{12}/(s_{11}s_{12})^{1/2} \tag{2.3}$$

$(i = 1, 2)$, and a maximal invariant parameter is

$$\Gamma = (\gamma_1, \gamma_2, \rho) \quad \text{with} \quad \gamma_i = \mu_i / \sigma_{ii}^{1/2} \quad \text{and} \quad \rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2} \tag{2.4}$$

$(i = 1, 2)$. For notation, let

$$\begin{cases} v_i = t_i/(1 + t_i^2)^{1/2}, \ (i = 1, 2), \ b = (t_1 t_2 + r)/[(1 + t_1^2)(1 + t_2^2)]^{1/2} \\ \tau = \gamma_1^2 - 2\rho\gamma_1\gamma_2 + \gamma_2^2, \ \delta_1 = \gamma_1 - \rho\gamma_2 \ \text{and} \ \delta_2 = \gamma_2 - \rho\gamma_1. \end{cases} \tag{2.5}$$

The following lemma is a generalization of Miller (1968) and Siddiqui (1967) and gives the distribution of (t_1, t_2, r).

Lemma 2.1. *The* pdf *of* $T(t_1, t_2, r)$ *is given by*

$$f(T|\Gamma) = h(T|\Gamma) g_1(t_1) g_2(t_2) g_3(r) \tag{2.6}$$

where, with $d_{n+i} = 2^{i/2}\Gamma\Big(\frac{n+i+1}{2}\Big)/\Gamma\Big(\frac{n+1}{2}\Big)$,

$$h(T|\Gamma) = \exp\Big(-\frac{\tau}{2}\Big)(1-\rho^2)^{-(n+1)/2}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}\frac{(v_1\delta_1)^j(v_2\delta_2)^k(b\rho)^\ell}{j!k!\ell!}d_{n+j+\ell}d_{n+k+\ell} \tag{2.7}$$

$$g_i(t_i) = \Big[\Gamma\Big(\frac{n+1}{2}\Big)/\Gamma\Big(\frac{n}{2}\Big)\pi^{1/2}\Big](1+t_i^2)^{-(n+1)/2} \quad (i=1,2) \tag{2.8}$$

$$g_3(r) = \Big[\Gamma\Big(\frac{n}{2}\Big)/\Gamma\Big(\frac{n-1}{2}\Big)\pi^{1/2}\Big](1-r^2)^{(n-3)/2}. \tag{2.9}$$

Proof. Let P_Γ^T be the distribution of T under Γ. Then by Wijsman's Theorem (1967), the pdf of T with respect to P_0^T, evaluated at $T \equiv T(y,S)$, is given by $dP_\Gamma^T/dP_0^T = Q_\Gamma/Q_0$, where

$$Q_\Gamma = \int_G f_1(Ay|\mu,\Sigma)f_2(ASA'|\Sigma)\chi(A)\nu(dA) \tag{2.10}$$

where $f_1(y|\mu,\Sigma)$ and $f_2(S|\Sigma)$ are respectively the pdf's of y and S in (2.1), $\chi(A)$ is the inverse of the Jacobian of $(y,S) \to (Ay, ASA')$, i.e., $\chi(A) = a_1^4a_2^4$ and $\nu(dA) = (a_1a_2)^{-1}da_1da_2$ is an invariant measure on G. First using the invariance of ν and replacing a_i by $a_i/s_{ii}^{1/2}\sigma_{ii}^{1/2}$ $(i=1,2)$ in (2.10), Q_Γ is evaluated as

$$Q_\Gamma \propto c(\theta)\int_G \exp\Big[-\frac{1}{2}\text{tr}\Lambda A(tt'+R)A'\Big]\exp[\text{tr}\Lambda At\gamma']a_1^{n+1}a_2^{n+1}\nu(dA) \tag{2.11}$$

where $c(\theta) = |\Sigma|^{-(n+1)/2}\exp(-\frac{1}{2}\mu'\Sigma^{-1}\mu) = [\sigma_{11}\sigma_{22}(1-\rho^2)]^{-(n+1)/2}\exp(-\frac{\tau}{2})$, $t = (t_1,t_2)'$, $R = \begin{pmatrix}1 & r\\ r & 1\end{pmatrix}$ and $\Lambda = \begin{pmatrix}1 & -\rho\\ -\rho & 1\end{pmatrix}$. The constants, such as $|S|^{(n-3)/2}$, which are cancelled out with those of the denominator Q_0, are absorbed into the notation $\propto$. Next again replacing a_i by $a_i/(1+t_i^2)^{1/2}$ in (2.11) yields

$$Q_\Gamma \propto c(\theta)\int_0^\infty\int_0^\infty a_1^na_2^n\exp(a_1v_1\delta_1+a_2v_2\delta_2)\exp(a_1a_2b\rho)\exp\Big(-\frac{1}{2}(a_1^2+a_2^2)\Big)da_1da_2 \tag{2.12}$$

with v_i's, δ_i's and b from (2.5). Here, on expanding $\exp(a_iv_i\delta_i)$'s and $\exp(a_1a_2b\rho)$, (2.12) becomes

$$Q_\Gamma \propto c(\theta)\sum_j\sum_k\sum_\ell\frac{(v_1\delta_1)^j(v_2\delta_2)^k(b\rho)^\ell}{j!k!\ell!}\iint a_1^{n+j+\ell}a_2^{n+k+\ell}\exp\Big(-\frac{1}{2}(a_1^2+a_2^2)\Big)da_1da_2. \tag{2.13}$$

Hence, using $\int_0^\infty a^m\exp(-(1/2)a^2)da = s^{(m-1)/2}\Gamma((m+1)/2)$ and taking the ratio of Q_Γ and Q_0, the expression $h(T|\Gamma)$ in (2.7) is obtained. On the other hand, t_1, t_2 and r are independent when $\Gamma = (\gamma_1, \gamma_2, \rho) = 0$ and the pdf's of these variables are given by (2.8) and (2.9) respectively. This implies $dP_0^T = g_1(t_1)g_2(t_2)g_3(r)dt_1dt_2dr$, and so $[Q_\Gamma/Q_0]dP_0^T = f(T|\Gamma)dt_1dt_2dr$, which completes the proof.

For the case $\mu = 0$, Siddiqui (1967) derived the exact and approximate pdf of (t_1, t_2, r). The pdf of $T = (t_1, t_2, r)$ with $\mu = 0$ is $f(t_1, t_2, r|0, 0, \rho)$ since $\mu = 0$ is equivalent to $\delta_1 = \delta_2 = 0$. The distribution of (t_1, t_2) in the following lemma is regarded as a two-dimensional noncentral t distribution since $t_i = y_i/s_{ii}^{1/2}$ $(i = 1, 2)$.

Lemma 2.2. *The* pdf *of* (t_1, t_2) *is given by*

$$f_0(t_1, t_2|\gamma_1, \gamma_2, \rho) = h_0(t_1, t_2|\gamma_1, \gamma_2, \rho)g_1(t_1)g_2(t_2) \tag{2.14}$$

where, with $e_q = \Gamma\left(\frac{q+1}{2}\right)\Gamma\left(\frac{n}{2}\right)/2\Gamma\left(\frac{q+n}{2}\right)\pi^{1/2}$,

$$\begin{aligned} &h_0(t_1, t_2|\gamma_1, \gamma_2, \rho) \\ &\quad = \exp\left(-\frac{\tau}{2}\right)(1-\rho^2)^{-(n+1)/2}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}\frac{(v_1\delta_1)^j(v_2\delta_2)^k\rho^\ell}{j!k!\ell!}d_{n+j+\ell}d_{n+k+\ell} \\ &\quad \times\sum_{q=0}^{\ell}[e_q + (-1)^q e_q]\binom{\ell}{q}(v_1v_2)^{\ell-q}[(1-v_1^2)(1-v_2^2)]^{q/2}. \end{aligned} \tag{2.15}$$

Proof. Expanding $b^\ell = [r(1-v_1^2)^{1/2}(1-v_2^2)^{1/2} + v_1v_2]^\ell$ as a binomial series and integrating it with respect to $g_3(r)$ yields the result.

Miller (1968) treated the multivariate t-distribution of (t_1, t_2) with $\rho = 0$ and some other associated distributions. It seems that the above two dimensional noncentral t-distribution has not been derived.

Lemma 2.2 implies that a scale invariant test is not similar in general since the distribution of the maximal invariant under H depends on the nuisance parameter ρ. Of course, such an invariant test as a test based on t_1 only is similar. A question to be posed here is whether the problem admits a group structure under which the class of invariant tests forms a class of similar tests. In particular, is there a nontrivial permutation invariant and invariantly similar test? This question is treated in the next section.

3. Invariance Analysis of the Problem

By the traditional invariance approach we study here, the invariant structure of the problem.

The problem stated in Section 1 may be viewed as the problem of testing

$$H: \ \mu = 0 \quad \text{vs} \quad K: \ \mu \geq 0, \ \mu \neq 0 \quad \text{with} \quad \mu = \sqrt{N}\eta \tag{3.1}$$

in the transformed model

$$X = \begin{pmatrix} y' \\ Z \end{pmatrix} : n \times p \sim N_{N \times p}\Big(\begin{pmatrix} \mu' \\ 0 \end{pmatrix}, I_N \otimes \Sigma\Big). \tag{3.2}$$

In terms of the sufficient statistic (y, S) with $S = Z'Z$, the model is

$$y \sim N(\mu, \Sigma) \quad \text{and} \quad S \sim W(\Sigma, N-1) \tag{3.3}$$

where $W(\Sigma, m)$ denotes the Wishart distribution with mean $m\Sigma$ and degrees of freedom m. Let $\Theta_H = \{(0, \Sigma) | \Sigma \in S(p)\} = \{0\} \times S(p)$ and $\Theta_K = \{(\mu, \Sigma) | \mu \geq 0, \Sigma \in S(p)\} = (R_+^p - \{0\}) \times S(p)$ where $R_+ = [0, \infty)$. As has been seen in Section 2, the problem is left invariant under the group

$$G_1 = A(p) \tag{3.4}$$

which acts on the sufficient statistic (y, S) by

$$g \circ (y, S) = (Ay, ASA') \quad \text{with } g = A \in G_1 \tag{3.5}$$

where $A(p)$ is the group of $p \times p$ diagonal matrices with positive diagonal elements. On the other hand, the problem is also left invariant under the group

$$G_2 = P(p), \tag{3.6}$$

which acts on (y, S) by $g \circ (y, S) = (Qy, QSQ')$ where $g = Q \in G_2$. The group defined by

$$G_3 = AP(p) \qquad \text{with} \tag{3.7}$$

$$AP(p) = \{B | B = B_1 B_2 \cdots B_m, B_i \in A(p) \text{ or } B_i \in P(p), i = 1, \ldots, m, m \in \mathbb{N}\} \tag{3.8}$$

leaves the problem (3.1) invariant, where G_3 acts on (y, S) by $g \circ (y, S) = (By, BSB')$ with $g = B \in AP(p)$ and it acts on Σ by $g \circ \Sigma = B\Sigma B'$. Clearly G_3 contains G_1 and G_2 as subgroups, and hence neither G_1 nor G_2 is maximal as a group leaving the problem (3.1) invariant.

Now define

$$\begin{cases} \tilde{T} = \tilde{T}(y, S) = (t_{(1)}, \ldots, t_{(p)} : \tilde{r}_{12}, \ldots, \tilde{r}_{(p-1)p}) \\ \tilde{\Gamma} = \tilde{\Gamma}(\mu, \Sigma) = (\gamma_{(1)}, \ldots, \gamma_{(p)} : \tilde{\rho}_{12}, \ldots, \tilde{\rho}_{(p-1)p}) \end{cases} \tag{3.9}$$

where $t_{(1)} \geq \cdots \geq t_{(p)}$, i.e. $(t_{(1)}, \ldots, t_{(p)})' = Q_t(t_1, \ldots, t_p)'$ (for $Q_t \in P(p)$) (or $\gamma_{(1)} \geq \cdots \geq \gamma_{(p)}$) are the ordered t_i's (or γ_i's) with $t_i = y_i / \sqrt{s_{ii}}$ (or $\gamma_i = \mu_i / \sqrt{\sigma_{ii}}$)

and $r_{ij} = s_{ij}/(s_{ii}s_{jj})^{1/2}$ (or $\rho_{ij} = \sigma_{ij}/(s_{ii}s_{jj})^{1/2}$) $(i \neq j)$ with $S = (s_{ij}) = Z'Z$, $(\tilde{r}_{ij}) = Q_t R Q_t'$ and $R = (r_{ij})$.

Lemma 3.1. *(1) $\tilde{T}$ is maximal invariant under G_3. (2) $\tilde{\Gamma}$ is a maximal invariant parameter under G_3.*

Proof. The proof is straightforward and therefore omitted.

This lemma shows that a G_3 invariant test is not similar in general and that there is no further dimensional reduction on the space of a maximal invariant parameter under G_1. In fact, since $\gamma_{(i)} = 0$ $(i = 1, \ldots, p)$ under H, the null distribution of $\tilde{T}$ depends on ρ_{ij}'s by (2) of Lemma 3.1 and $\dim G_3 = \dim G_1$. Further, even the test based on the critical region $t_{(1)} > c$ is not similar, though the test based on $t_1 > c$ is similar. To see this, consider the case $p = 2$. Then from Lemma 2.2, the pdf of $(t_{(1)}, t_{(2)})$ is obtained as $2f_0(t_{(1)}, t_{(2)}|\gamma_{(1)}, \gamma_{(2)}, \rho)$ with $t_{(1)} \geq t_{(2)}$, where $f_0(t_1, t_2|\gamma_1, \gamma_2, \rho)$ is given in (2.4). From this pdf, it can be directly observed that the marginal pdf of $t_{(1)}$ under $H : \gamma_{(1)} = \gamma_{(2)} = 0$ does depend on ρ.

To summarize, neither G_1 nor G_2 nor G_3 provide similar tests through invariance only. This is because the groups are too small to act on Θ_H transitively under H. In Section 2, we observed this for G_1 by deriving the pdf of a maximal invariant. For G_2, it is easy to see since the invariance under $P(p)$ does not reduce the dimension of the original parameter space. For G_3, it was discussed above.

A question we may ask here is whether G_3 is a maximal group as a group leaving the problem invariant. The following lemma answers it positively.

Lemma 3.2. *The group G_3 is a maximal group leaving the model $M(\Theta_K) = \{N(\mu, \Sigma)|(\mu, \Sigma) \in \Theta_K\}$ invariant.*

Proof. By Nabeya and Kariya (1986), any measurable transformation $h(y, S)$ which preserves the model $M(\Theta_K) = \{N(\mu, \Sigma) \times W(\Sigma, N-1)|(\mu, \Sigma) \in \Theta_K\}$ must be of the form $h(y, S) = (Cy, CSC')$ a.e. for some $C \in G\ell(p)$. Hence the group of transformations preserving $M(\Theta_K)$ is regarded as

$$J_p = \{C \in G\ell(p)|C\mu \geq 0 \text{ and } C^{-1}\mu \geq 0 \text{ for all } \mu \geq 0\}.$$

We shall show $J_p = G_{3p}$ where the suffix p denotes the dimension of μ. Since $G_{3p} \subset J_p$, we show $J_p \subset G_{3p}$ by mathematical induction on p. It is clear for the case $p = 1$. Assuming it for the case p, consider the case $p+1$. Let $\tilde{C} \in J \equiv J_{p+1}$. Since $\tilde{C} \in J_{p+1}$ is equivalent to $Q\tilde{C}Q' \in J_{p+1}$ for any given $Q \in P(p+1)$, take

$Q \in P(p+1)$ such that $C_{22} \neq 0$. Here,

$$Q\tilde{C}Q' \equiv C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \quad C^{-1} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \quad \text{and} \quad \mu = \begin{pmatrix} \mu_p \\ \mu_{p+1} \end{pmatrix}$$

where $C_{11} : p \times p$, $D_{11} : p \times p$ and $\mu_p \in R^p$. Note $C_{22} \neq 0$ implies $|C_{11.2}| \neq 0$ with $C_{11.2} = C_{11} - C_{12}C_{22}^{-1}C_{21}$. Now the statement that $C\mu \geq 0$ and $C^{-1}\mu \geq 0$ for all $\mu \geq 0$ is equivalent to the statement

$$C_{11}\mu_p + C_{12}\mu_{p+1} \geq 0, \quad C_{21}\mu_p + C_{22}\mu_{p+1} \geq 0 \tag{3.10}$$

$$D_{11}\mu_p + D_{12}\mu_{p+1} \geq 0, \quad D_{21}\mu_p + D_{22}\mu_{p+1} \geq 0 \tag{3.11}$$

for all $\mu \geq 0$. It follows from (3.10) that

$$C_{12} \geq 0,\ C_{21} \geq 0,\ C_{22} > 0 \text{ and } C_{11}\mu_p \geq 0 \text{ for all } \mu_p \geq 0$$

and it follows from (3.11) that

$$D_{12} \geq 0,\ D_{21} \geq 0,\ D_{22} \geq 0 \text{ and } D_{11}\mu_p \geq 0 \text{ for all } \mu_p \geq 0.$$

Also from $CC^{-1} = I$, it follows that

$$D_{22} = C_{22}^{-1} + C_{22}^{-1}C_{21}C_{11.2}^{-1}C_{12}C_{22}^{-1} > 0, \quad D_{12} = -C_{11.2}^{-1}C_{12}C_{22}^{-1} \geq 0,$$
$$D_{11} = (C_{11} - C_{12}C_{22}^{-1}C_{21})^{-1}, \quad D_{21} = -C_{22}^{-1}C_{21}C_{11.2}^{-1} \geq 0.$$

Hence $D_{21}D_{11}\mu_p = -C_{22}^{-1}C_{21}\mu_p \geq 0$ for all μ_p, implying $C_{21} \leq 0$ by $C_{22} > 0$. Therefore $C_{21} = 0$, and so $D_{11} = C_{11}^{-1}$. Thus, it follows that $C_{11}\mu_p \geq 0$ and $C_{11}^{-1}\mu_p \geq 0$ for all $\mu_p \geq 0$. This implies $C_{11} \in G_{3p}$ by the assumption of induction. Hence the elements of C_{11} and C_{11}^{-1} are nonnegative, and so $C_{12} \leq 0$ from $D_{12} \geq 0$, implying $C_{12} = 0$. Consequently we have proved that $\tilde{C} = Q'CQ$ with $C_{12} = 0$, $C_{21} = 0$, $C_{22} > 0$ and $C_{11} \in G_{3p}$. Therefore $\tilde{C} \in G_{3(p+1)}$ by the definition of G_3 in (3.7), completing the proof.

This lemma implies that no larger group including G_3 leaves the problem invariant, and hence no invariantly similar tests will be provided through the traditional invariance approach. In other words, the traditional invariance approach does not lead us to conclude that a class of similar tests is the class of invariant tests in this problem. An alternative approach to overcome such a situation will be discussed in the next section.

4. *H–K* Invariance Approach

In Section 3, it has been observed that the traditional invariance approach provides no nontrivial similar tests because the group leaving the problem invariant is too small. The fact that the group G_3 in (3.9) is maximal, implies that there is no way to obtain nontrivial similar tests through the traditional invariance approach. Therefore, if similarity is considered as an important property for a test, the traditional invariance approach should be modified to provide invariantly similar tests. An attempt to do so is presented in this section.

To describe our modification of the traditional invariance approach, consider a general testing problem:

$$H : \theta \in \Theta_H \quad \text{vs} \quad K : \theta \in \Theta_K \quad (\Theta_H \cap \Theta_K = \emptyset) \tag{4.1}$$

where X is a sample space in R^n and $M(\Theta) = \{P_\theta : \theta \in \Theta\}$ ($\Theta = \Theta_H \cup \Theta_K \subset R^m$) is a set of probability measures on X indexed by $\theta \in \Theta$. As is well known, the model $M(\Theta)$ is said to be invariant under a group G of measurable transformations of X if $gM(\Theta) = M(\Theta)$ i.e. for any measurable set A, $gP_\theta(A) \equiv P_\theta(g^{-1}(A))$ and $gP_\theta \in M(\Theta)$ for all $g \in G$. The testing problem (4.1) is said to be invariant under G if both the null model $M(\Theta_H)$ and the nonnull model $M(\Theta_K)$ are invariant under G. Hence in the traditional invariance approach, we are required to find a (maximal) group leaving both models invariant. Sometimes we face a case where the group found becomes too small to provide invariantly similar tests. An approach for such a case is to deal separately with the groups leaving each model invariant. To state it, let

$$\mathbf{H} = \{H \in \mathbf{C} | H \text{ leaves } M(\Theta_H) \text{ invariant}\} \tag{4.2}$$

$$\mathbf{H}_S = \{H \in \mathbf{H} | H \text{ acts transitively on } \Theta_H\} \tag{4.3}$$

and

$$\mathcal{K} = \{K \in \mathbf{C} | K \text{ leaves } M(\Theta_K) \text{ invariant}\}, \tag{4.4}$$

where $\mathbf{C}$ is the set of groups acting (measurably) on X. Here we assume that $\Theta_H \subset \bar{\Theta}_K$, i.e., Θ_H is in the boundary of Θ_K and that $\mathbf{H} \supset \mathbf{H}_S \supset \mathcal{K}$, which is often the case in applications, especially in nested problems. Our approach here is to first choose a group K from $\mathcal{K}$ which may be maximal (though it is not necessarily required), and, secondly, to choose a minimal group H_m from $\mathbf{H}_S$ that includes K as a subgroup (if any). Thirdly, in the class of H_m invariant tests, an optimal or appropriate test is to be chosen. Of course, the transitivity of H ($\in \mathbf{H}_S$) on Θ_H guarantees the similarity of an H invariant test. Here it is noted that though an H_m invariant test is K_M invariant, it is supposedly a test for the problem

$$H : \theta \in \Theta_H \quad \text{vs} \quad \tilde{K} : \theta \in \tilde{\Theta}_K \equiv \{h\theta | \theta \in \Theta_K, h \in H_m\}. \tag{4.5}$$

Hence, the minimality required in our approach means that $\tilde{\Theta}_K$ should be close to Θ_K.

The minimality of a group H in H_S is checked as follows. Since H acts on Θ_H as an induced group, let $H_\theta = \{h \in H | h\theta = \theta\}$ be the isotropy group of θ. When $H_\theta = \{e\}$ for all $\theta \in \Theta_H$ with e the identity of H, H is said to act freely on Θ_H.

Lemma 4.1. *If the action of H is free as well as transitive, then H is minimal in H_S.*

Proof. Suppose that there exists a subgroup $H_0 \subset H$ such that the action of H_0 on Θ_H is transitive. Then to claim $H_0 = H$, take any $h \in H$ and $\theta \in \Theta_H$. Since H_0 is transitive, there exists an $h_0 \in H_0$ such that $h_0\theta = h\theta$. Hence, $h^{-1}h_0 \in H_\theta$, but $H_\theta = \{e\}$ which implies $h = h_0 \in H_0$. This completes the proof.

Finally, we shall define a measure for the nonsimilarity of the invariant problem (4.1) by the difference of the dimensions of the minimal group H_m and the (maximal) group K:

$$d = \dim H_m - \dim K \quad (\geq 0) \tag{4.6}$$

where H_m is assumed to be a matrix group. If $H_m = K$, the problem is invariant in the traditional sense, in which case $d = 0$. Though $d = 0$ does not necessarily imply $H_m = K$ in general, the "extra" part $H_m - K$ in case of $d = 0$ does not contribute to the reduction of the dimension of a maximal invariant under K. In most applications, $d = 0$ implies $H_m = K$. When $0 < d < \infty$, then d corresponds to the difference of the dimensions of the maximal invariants under H_m and K, and it also corresponds to the difference of dimensions of maximal invariant parameters when H_m and K are the induced groups.

It is noted that we do not necessarily require the maximality of K in $\mathcal{K}$. The reason will be made clear later.

5. The H–K Approach to the Problem

We apply the formulation made in Section 4 to the problem (3.1) with $K = G_1$ and $K = G_3$ as a group leaving the problem invariant, where $K \in \mathcal{K}$ in (4.4). First consider the group G_1 in (3.3) as a group K_1 in $\mathcal{K}$ of (4.4) leaving the nonnull model $M(\Theta_K)$ invariant. As has been seen in Section 2, this group $K_1 \equiv G_1$ does not act transitively on Θ_H and a class of similar tests is not provided as the class of invariant tests under K_1. Hence to find a minimal group H_1 in H_S of (4.3) which includes K_1 as a subgroup, let

$$H_1 = T(p) \tag{5.1}$$

where $T(p)$ is the group of $p \times p$ nonsingular lower triangular matrices with positive diagonal elements. H_1 acts on (y, S) by $h \circ (y, S) = (Uy, USU')$ and on Σ by $h \circ \Sigma = U\Sigma U'$ where $h = U \in T(p)$.

Lemma 5.1. *H_1 is a minimal group H_m in* H *of* (4.3) *which includes $K_1 = G_1$.*

Proof. Since $K_1 \subset H_1$, by Lemma 4.1, it suffices to show that H_1 acts transitively and freely on Θ_H. The transitivity is clear since $\Sigma \in S(p)$ is uniquely decomposed as $\Sigma = \Psi\Psi'$ with $\Psi \in T(p)$. The freeness also follows. In fact, $U\Sigma U' = \Sigma$ implies $U\Psi = \Psi$ by the uniqueness of the decomposition, which in turn implies $U = I$ i.e., the isotropy group $H_{1\Sigma}$ equals $\{I\}$ for any $\Sigma \in S(p)$. This completes the proof.

The degree of the nonsimilarity of the invariant problem (3.1), defined by (4.6) is

$$d = [p(p+1)/2] - p = p(p-1)/2, \tag{5.2}$$

which corresponds to the dimension of the nuisance parameter of a maximal invariant parameter under K_1 when H holds. When $p = 2$, $d = 1$ as has been observed in Section 2.

As in (4.5), an H_1-invariant test presumably tests $H : \theta \in \Theta_H$ vs

$$\begin{aligned} \tilde{K} : \theta \in \tilde{\Theta}_K &= \{(U\mu, U\Sigma U')|\mu \geq 0, \Sigma \in S(p), U \in T(p)\} \\ &= \{x \in R^p | x_1 \geq 0\} \times S(p) = (R_+ \times R^{p-1}) \times S(p), \end{aligned} \tag{5.3}$$

which is different from $\Theta_K = R_+^p \times S(p)$.

A maximal invariant under H_1 is

$$w \equiv (w_i) = W^{-1/2} y \quad \text{with} \quad W = Z'Z + yy' \quad \text{and} \quad W^{-1/2} \in T(p) \tag{5.4}$$

and a maximal invariant parameter is

$$\beta \equiv (\beta_i) = \Sigma^{-1/2}\mu \quad \text{with} \quad \Sigma^{-1/2} \in T(p). \tag{5.5}$$

Theorem 5.1. (1) *The* pdf *of w under H is $f_0(w) = c_0(1 - w'w)^{(N-p-2)/2}$.* (2) *The* pdf *of w under K is*

$$f(w|\beta) = \exp\Big(-\frac{1}{2}\beta'\beta\Big) \exp\Big(\frac{1}{2}\sum_{j=1}^{p} w_j^2 \sum_{i=j+1}^{p} \beta_i^2\Big)\Big[\prod_{i=1}^{p} F_i(w_i\beta_i)\Big] f_0(w) \tag{5.6}$$

where

$$F_i(w_i\beta_i) = \sum_{k=0}^{\infty} (\sqrt{2}\beta_i w_i)^k \Gamma((a_i + k)/2)/k!\Gamma(a_i) \quad \textit{with} \quad a_i = N = i + 1. \tag{5.7}$$

Proof. (1) The pdf of (y, S) with $S = Z'Z$ is $c|\Sigma|^{-N/2}\exp[-\frac{1}{2}\mathrm{tr}\Sigma^{-1}(S+yy')]\times |S|^{\alpha}dydS$ where $\alpha = (N-p-2)/2$. In this pdf, transforming S into W and then y into w yields the pdf of (w, W):

$$c|\Sigma|^{-N/2}\exp\Big(-\frac{1}{2}\mathrm{tr}\Sigma^{-1}W\Big)|W|^{\alpha+1/2}|I-ww'|^{\alpha}dwdW \quad (w'w \le 1).$$

Integrating W out gives the result. (2) Let $f(X|\mu,\Sigma)$ be the pdf of X in (3.2) and let P_β^w be the distribution of w under β. Then by Wijsman's Theorem (1967), $dP_\beta^w/dP_0^w = H(X|\mu,\Sigma)/H(X|0,\Sigma)$ with

$$H(X|\mu,\Sigma) = \int_{H_m} f(g\circ X|\mu,\Sigma)\chi(g)\nu(dg)$$

where $\chi(g) = |UU'|^{N/2}$ for $g = (R, U) \in H_m$, $\nu(dg) = \nu_1(dR)\nu_2(dU)$, $\nu_1(dR)$ is the invariant probability measure on $O(N-1)$ and $\nu_2(dU) = \prod_{i=1}^p u_{ii}^{-1}dU$ is an invariant measure on $T(p)$. Since $H(X|\mu,\Sigma)$ is proportional to

$$\exp\Big(-\frac{1}{2}\mu'\Sigma^{-1}\mu\Big)\int_{T(p)}\exp(\mu'\Sigma^{-1}Uy)\exp\Big(-\frac{1}{2}\mathrm{tr}U'\Sigma^{-1}UW\Big)|UU'|^{N/2}\nu_2(dU),$$

transforming U into $\Sigma^{-1/2}UW^{-1/2}$ yields

$$H(X|\mu,\Sigma) \propto \exp\Big(-\frac{1}{2}\beta'\beta\Big)\int_{T(p)}\exp(\beta'Uw)\exp\Big(-\frac{1}{2}\mathrm{tr}UU'\Big)|UU'|^{N/2}\nu_2(dU).$$

Therefore $dP_\beta^w/dP_0^w = \exp(-\frac{1}{2}\beta'\beta)E^U[\exp(\beta'Uw)]$. Since u_{ij}'s are independently distributed with $u_{ij} \sim N(0,1)$ for $i \ne j$ and $u_{ii}^2 \sim$ Gamma$((N-i+1)/2)$ with pdf $v^{a-1}\exp(-a/2)/\Gamma(a)2^a$ where $a = (N-i+1)/2$, and since $E^U[\exp(\beta'Uw)] = E[\exp(\sum_{j=1}^p\sum_{i=j+1}^p w_j\beta_i u_{ij})\exp(\sum_{i=1}^p u_{ii}w_i\beta_i)]$, the result follows from $E[\exp(w_j\beta_i u_{ij})] = \exp(-\frac{1}{2}w_j^2\beta_i^2)$ and $E[\exp(u_{ii}w_i\beta_i)] = F_i(w_i\beta_i)$.

Next we consider the case where a group K leaving the problem (3.1) invariant is chosen to be the maximal group $G_3 = AP(p)$. As a group H which acts transitively on Θ_H and includes K, we choose

$$H_3 = G\ell(p) \tag{5.8}$$

which acts on (y, S) by $g\circ(y,S) = (Cy, CSC')$ and on Θ by $g\circ(\mu,\Sigma) = (C\mu, C\Sigma C')$ where $g = C \in H_3$.

Lemma 5.2. *H_3 is a minimal group H_m which acts transitively on Θ_H and includes $K_3 = G_3$ as a subgroup.*

Proof. Clearly H_3 acts transitively on Θ_H and it includes K_3. To show the minimality note that $K_3 = G_3$ contains $K_1 = G_1 = A(p)$ and that H_1 is the minimal group which acts transitively on Θ_H and includes K_1. This implies that the minimal group H_m contains $H_1 = T(p)$. Further, since K_3 contains $P(p)$, H_m contains $P(p)$ as well as $T(p)$, which implies that H_m contains $U(p)$, the group of $p \times p$ upper triangular matrices with positive diagonal elements. On the other hand, any matrix $C \in G\ell(p)$ is expressed as a product elementary matrices, $C = E_1 E_2 \cdots E_m$, where E_i's are elementary matrices. Since an elementary matrix belongs to one of $T(p)$, $U(p)$ and $P(p)$ (see, e.g., Rao (1973)), and since H_m contains $T(p)$, $U(p)$ and $P(p)$, $H_m = H_3$ follows. This completes the proof.

As is well known, a maximal invariant under $H_3 = G\ell(p)$ is the well known Hotelling T^2 statistic;

$$u \equiv y'S^{-1}y$$

which when multiplied by $[(N-p)/(N-1)p]$ is distributed as F with degrees of freedom $(p, N-p)$. An H_3-invariant test presumably tests $H : \theta \in \Theta_H$ vs

$$\begin{aligned} K^* : \theta \in \Theta_K^* &= \{(C\mu, C\Sigma C') | C \in G\ell(p), \mu \neq 0, \Sigma \in S(p)\} \\ &= (R^p - \{0\}) \times S(p), \end{aligned}$$

which is known as the Hotelling T^2 problem, and the test with critical region $u > c$ is UMPI (uniformly most powerful invariant) for H vs K^*. Therefore, by Lemma 5.1 and Lemma 5.2 we conclude

Theorem 5.1. *The Hotelling T^2 test statistic is the unique invariantly similar statistic which is coodinatewise scale invariant and permutation invariant.*

6. Non-Bayes Property of Hotelling T^2 Test

In Section 5, we observed that the requirement for a test to be scale invariant coordinatewise, symmetric coordinatewise and similar eventually leads us to the Hotelling T^2 statistic as a maximal invariant under $G_3 = G\ell(p)$, because $G_1 = U(p)$ and $G_2 = P(p)$ generate G_3. In this section, we study a property of the Hotelling T^2 test in the problem.

Theorem 6.1. *The Hotelling T^2 test cannot be a proper Bayes test.*

Proof. Let $\delta = \Sigma^{-1}\mu$ and $\Phi = \Sigma^{-1}$. Then any Bayes test is of the form; reject H if

$$b(y, W) \equiv \frac{\int_\partial \exp(-\frac{1}{2}\mathrm{tr}\Phi W + y'\delta)d\pi}{\int_\partial \exp(-\frac{1}{2}\mathrm{tr}\Phi W)d\pi} > c \tag{6.1}$$

with respect to some finite measure π over ∂ where

$$\partial = \{(\delta, \Phi) | \Phi^{-1}\delta \geq 0, \Phi \in S(p)\} \tag{6.2}$$

and $W = S + yy'$ ($\geq yy'$) with $S = Z'Z$ (see (3.2)). Note that $b(y, W)$ is a continuous function on $R^p \times S(p)$. In fact it is analytic. Now suppose that the T^2 test is Bayes. Then for some finite measure π and for some increasing function f, the following condition must be satisfied

$$b(y, W) = f(y'W^{-1}y). \tag{6.3}$$

This is because the set of analytic sets $\{\{(y, W)|b(y, W) \leq c\},\ c \in R\}$ must generate the set $\{\{(y, W)|y'W^{-1}y \leq d\},\ d \in R\}$. Note that $y'W^{-1}y$ is in one-one correspondence with $T^2 = y'S^{-1}y$. It follows from (6.3) that $b(Ay, AWA') = b(y, W)$ for all $A \in G\ell(p)$. Hence taking $A = W^{-1/2}$

$$b(y, W) = \int_{\partial} \exp\Big(-\frac{1}{2}\mathrm{tr}\Phi + z'\delta\Big)d\pi \Big/ \int_{\partial} \exp\Big(-\frac{1}{2}\mathrm{tr}\Phi\Big)d\pi = f(z'z) \tag{6.4}$$

where $z = W^{-1/2}y$. Define a probability measure on ∂ by

$$d\nu = \exp\Big(-\frac{1}{2}\mathrm{tr}\Phi\Big)d\pi \Big/ \int_{\partial} \exp\Big(-\frac{1}{2}\mathrm{tr}\Phi\Big)d\pi. \tag{6.5}$$

Then (6.4) implies

$$E_{\nu}[\exp(z'\delta)] = f(z'z) \quad (z \in R^p). \tag{6.6}$$

The left hand side of (6.6) is the moment generating function of δ, while the right hand side shows that the distribution of δ must be spherical. However, the domain of δ is restricted to $\partial = \{\delta|\Phi^{-1}\delta \geq 0,\ \Phi \in S(p)\}$. Since $R^p - \partial_p$ is shown to contain an open set, the distribution of δ cannot be spherical, giving a contradiction. This completes the proof.

The proof of Theorem 6.1 shows that T^2 cannot be generalized Bayes either. However it is conceivable that T^2 could still be a weak limit of a sequence of Bayes tests. The weak limits form an essentially complete class.

To complement the proof of Theorem 6.1, let

$$\partial_p = \{\delta \in R^p|\delta = \Sigma^{-1}\mu,\ \mu \geq 0,\ \Sigma \in S(p)\} \tag{6.7}$$

and

$$Q_p = R^p - Q_p^* \quad \text{with} \quad Q_p^* = \{x \in R^p|x \leq 0\} - \{0\}. \tag{6.8}$$

Lemma 6.1. $\partial_p = Q_p$.

Proof. To show $\partial_p \subset Q_p$, take $\delta = \Sigma^{-1}\mu \in \partial_p$. Then $\mu'\delta = \mu'\Sigma^{-1}\mu \geq 0$. Since $\mu \geq 0$, this implies $\delta \notin Q_p^*$ and hence $\delta \in Q_p$. To show $Q_p \subset \partial_p$, we use mathematical induction on p. When $p = 1$, clearly $Q_1 = \partial_1$. Assume if for the case p and consider the case $p + 1$. Let

$$\Sigma_{p+1}^{-1} = \begin{bmatrix} \Lambda_p & \alpha_p \\ \alpha_p & \tau \end{bmatrix}, \quad \mu_{p+1} = \begin{bmatrix} \mu_p \\ \beta \end{bmatrix} \quad \text{and} \quad \delta_{p+1} = \begin{bmatrix} \delta_p \\ \gamma \end{bmatrix}. \tag{6.9}$$

We want to show that for any given $\delta_{p+1} \in Q_{p+1}$, there exist $\mu_{p+1} \geq 0$, Λ_p, α_p and τ such that $\Sigma_{p+1}^{-1}\mu_{p+1} = \delta_{p+1}$. If $\delta_{p+1} = 0$, by taking $\mu_{p+1} = 0$, we note that any choice of Λ_p, α_p and τ will do. Suppose $\delta_{p+1} \neq 0$. Then $\Sigma_{p+1}^{-1}\mu_{p+1} = \delta_{p+1}$ is equivalent to

$$\Lambda_p\mu_p + \alpha_p\beta = \delta_p \quad \text{and} \quad \alpha_p'\mu_p + \tau\beta = \gamma. \tag{6.10}$$

In δ_{p+1} of (6.9), with $\gamma \geq 0$ and setting $\alpha_p = 0$ in (6.10) gives $\beta = \gamma/\tau \geq 0$ and $\Lambda_p\mu_p = \delta_p$. But by the assumption, $\Lambda_p\mu_p = \delta_p$ implies $\mu_p \geq 0$. Hence if $\gamma \geq 0$, there exists $\mu_{p+1} \geq 0$ such that $\delta_{p+1} = \Sigma_{p+1}^{-1}\mu_{p+1}$. Now suppose $\gamma < 0$. This implies $\delta_p \in Q_p$ since $\delta_{p+1} \in Q_{p+1}$. Choose $\beta = 0$ so that (6.10) becomes

$$\Lambda_p\mu_p = \delta_p \quad \text{and} \quad \alpha_p'\mu_p = \gamma. \tag{6.11}$$

We need to show that there exist Λ_p, α_p, τ and μ_p satisfying (6.11) and $\Sigma_{p+1}^{-1} > 0$. Since

$$\Sigma_{p+1}^{-1} = \begin{bmatrix} I & -\alpha_p/\tau \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \Lambda_p - \alpha_p\alpha_p'/\tau & 0 \\ 0 & \tau \end{bmatrix} \begin{bmatrix} I & 0 \\ -\alpha_p'/\tau & 1 \end{bmatrix}^{-1}$$

given any $\alpha_p \in R^p$ and $\Lambda_p \in S(p)$, there exists a $\tau > 0$ such that $\Sigma_{p+1}^{-1} > 0$. Hence by the assumption of induction, there exist $\Lambda_p \in S(p)$ and $\mu_p \geq 0$ such that $\Lambda_p\mu_p = \delta_p$ since $\delta_p \in Q_p$. Also for such μ_p, choose any $\alpha_p \in R^p$ such that $\alpha_p'\mu_p = \gamma < 0$. Therefore, given any $\delta_{p+1} \in Q_{p+1}$, we find $\mu_{p+1} \geq 0$ and $\Sigma_{p+1}^{-1} \in S(p+1)$ such that $\Sigma_{p+1}^{-1}\mu_{p+1} = \delta_{p+1}$, implying $\delta_{p+1} \in \partial_{p+1}$. This completes the proof.

Acknowledgement

Professor Cohen's research is supported by NSF grant #8822622.

References

Eaton, M. L. (1970). A complete class theorem for multidimensional one-sided alternatives. *Ann. Math. Statist.* **41**, 1884–1888.

Marden, J. I. (1982). Minimal complete classes of tests of hypotheses with multivariate one-sided alternatives. *Ann. Statist.* **10**, 962–970.

Miller, K. S. (1968). Some multivariate t-distributions. *Ann. Math. Statist.* **39**, 1605–1609.

Perlman, M. D. (1969). One-sided testing problems in multivariate analysis. *Ann. Math. Statist.* **40**, 549–567.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd edition. John Wiley, New York.

Siddiqui, M. M. (1967). A bivariate t distribution. *Ann. Math. Statist.* **38**, 162–166.

Tang, Dei-In (1990). Tests which are uniformly more powerful than the likelihood ratio tests for one-sided problems in multivariate analysis. Unpublished manuscript.

Wijsman, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Proc. 5th Berkeley Symp. Math. Statist. Probab.* **1**, 389-400. University of California, Berkeley.

Institute of Economic Research, Hitotsubashi University, Kunitachi, Tokyo 186, Japan.
Department of Statistics, Hill Center, Rutgers University, New Brunswick, NJ 08903, U.S.A.

(Received May 1990; accepted June 1991)