TESTING FOR INDEPENDENCE WITH ADDITIONAL INFORMATION

by

Morris L. Eaton[1] and Takeaki Kariya

Technical Report No. 238

University of Minnesota

Minneapolis, Minnesota

---

[1] This research was partially supported by National Science Foundation Grant NSF-GP-34482.

$$(2.30) \qquad \Pi(\phi,\beta) = \int \phi \, dP_{\beta}^{T} = \int \phi R dP_{0}^{T} = \mathcal{E}_{0} \, \phi R$$

$$= \mathcal{E}_{0} \, \phi[1 + \frac{n(n+m)}{2p_1p_2} \operatorname{tr}\bar{g}_1 x' \, \bar{g}_2'\bar{g}_2 x \, \bar{g}_1' \operatorname{tr} \beta\beta'$$

$$- \frac{(n+m)}{2p_2} \operatorname{tr} \bar{g}_2 u'u \, \bar{g}_2 \operatorname{tr} \beta\beta' + o(\operatorname{tr} \beta\beta')]$$

$$= \alpha + (\mathcal{E}_0 \, B(\phi)) \operatorname{tr} \beta\beta' + o(\operatorname{tr} \beta\beta')$$

where the remainder term is uniform in ϕ, and $B(\phi)$ is given by (2.4). The last equality in (2.30) follows by noting that

$$\operatorname{tr} \bar{g}_2 u'u \, \bar{g}_2' = \operatorname{tr} S_{22}(S_{22} + V)^{-1} \quad \text{and}$$

$$\operatorname{tr} \bar{g}_1 x' \, \bar{g}_2'\bar{g}_2 x \, \bar{g}_1' = \operatorname{tr} S_{11}^{-1} \, S_{12}(S_{22} + V)^{-1} \, S_{21} \, .$$

This completes the proof of Theorem 2.1.

$$(2.22)\qquad IV = \int_{\Theta(p_2)} \Big[\sum_{j=2}^{\infty} \frac{(-F)^j}{j!} + \sum_{j=2}^{\infty} \frac{E^{2j}}{(2j)!} + \sum_{\ell=1}^{\infty} \sum_{j=1}^{\infty} \frac{E^{2j}(-F)^{\ell}}{(2j)!\ \ell!} \Big] \mu_2^{(2)}(dk_2)$$

$$\equiv V + VI + VII \ .$$

Recall that for matrices $a: r_1 \times r_2$ and $b: r_2 \times r_1$, $|\mathrm{tr}\ ab| \le (\mathrm{tr}\ aa')^{\frac{1}{2}}(\mathrm{tr}\ bb')^{\frac{1}{2}}$ and $\mathrm{tr}(aa')^2 \le (\mathrm{tr}\ aa')^2$. Using these inequalities and arguing as in Schwartz (1967) yields

$$(2.23)\qquad \begin{cases} |-F|^j & \le (2p\ \mathrm{tr}\beta\beta')^j(\frac{1}{4}\ \mathrm{tr}\ h_2h_2')^j \\ E^{2j} & \le (4p\ \mathrm{tr}\ \beta\beta')^j(\frac{\mathrm{tr}\ h_1h_1' + \mathrm{tr}\ h_2h_2'}{4})^{2j} \end{cases}$$

where we have used the inequalities $\mathrm{tr}\ \bar{g}_2uu'\bar{g}_2' \le p$ and $\mathrm{tr}\ \bar{g}_1x'\bar{g}_2'\bar{g}_2\ x\bar{g}_1' \le p$. Thus for $\mathrm{tr}\ \beta\beta' < \frac{1}{4p}$,

$$(2.24)\qquad |V| \le \sum_{j=2}^{\infty}(j!)^{-1}[2p\ \mathrm{tr}\ \beta\beta']^2[\tfrac{1}{4}\ \mathrm{tr}\ h_2h_2']^j$$

$$\le [2p\ \mathrm{tr}\ \beta\beta']^2\ \exp[\tfrac{1}{4}\ \mathrm{tr}\ h_2h_2'] \ ,$$

$$(2.25)\qquad |VI| \le \sum_{j=2}^{\infty}[(2_j)!]^{-1}(4p\ \mathrm{tr}\ \beta\beta')^j[\frac{\mathrm{tr}\ h_1h_1' + \mathrm{tr}\ h_2h_2'}{4}]^{2j}$$

$$\le (4p\ \mathrm{tr}\ \beta\beta')^2\ \exp[\frac{\mathrm{tr}\ h_1h_1' + \mathrm{tr}\ h_2h_2'}{4}]$$

and

$$(2.26)\qquad |VII| \le \sum_{j=1}^{\infty}\sum_{\ell=1}^{\infty}[(2j)!\ \ell!]^{-1}[4p\ \mathrm{tr}\ \beta\beta']^j[4p\ \mathrm{tr}\ \beta\beta']^{\ell}$$

$$[\tfrac{1}{4}(\mathrm{tr}\ h_1h_1' + \mathrm{tr}\ h_2h_2')]^{2j}[\tfrac{1}{8}\ \mathrm{tr}\ h_2h_2']^{\ell}$$

$$\le [4p\ \mathrm{tr}\ \beta\beta']^2\ \exp[\tfrac{1}{4}\ \mathrm{tr}\ h_1h_1' + \tfrac{3}{8}\ \mathrm{tr}\ h_2h_2'] \ .$$

write $g_i = h_i k_i$ where $h_i \in G_T^+(p_i)$, $k_i \in \mathcal{O}(p_i)$ and let $\mu_1^{(i)}$ and $\mu_2^{(i)}$ be left invariant measures on $G_T^+(p_i)$ and $\mathcal{O}(p_i)$ respectively for $i = 1,2$. (For convenience we take invariant measure on $\mathcal{O}(r)$ to be invariant probability measure.) Then N in (2.12) can be written as

$$N = \int\int_{G^*}\int\int \exp(E-F)\Big[\prod_{i=1}^{2} |h_i h_i'|^{\frac{n(i)}{2}} \exp(-\tfrac{1}{2} \operatorname{tr} h_i h_i')\Big] \prod_{i=1}^{2} \mu_1^{(i)}(dh_i)\mu_2^{(i)}(dk_i) \tag{2.13}$$

where $G^* = G_T^+(p_1) \times G_T^+(p_2) \times \mathcal{O}(p_1) \times \mathcal{O}(p_2)$ and

$$\begin{cases} E = \operatorname{tr} \beta' h_2 k_2 \bar{g}_2\, x\, \bar{g}_1' k_1' h_1' \\ F = \tfrac{1}{2} \operatorname{tr} \beta\beta' h_2 k_2\, \bar{g}_2 (u'u)\, \bar{g}_2' k_2' h_2' \,. \end{cases} \tag{2.14}$$

Similarly, the denominator in (2.10) can be written as $D = D_1 D_2$ where

$$D_i = \int |h_i h_i'|^{\frac{n(i)}{2}} \exp[-\tfrac{1}{2} \operatorname{tr} h_i h_i'] \mu_1^{(i)}(dh_i)\mu_2^{(i)}(dk_i) \tag{2.15}$$

for $i = 1,2$. Defining $H_i(h_i)$, $i = 1,2$, by

$$H_i(h_i) = \frac{1}{D_i} |h_i h_i'|^{\frac{n(i)}{2}} \exp[-\tfrac{1}{2} \operatorname{tr} h_i h_i'] \,, \tag{2.16}$$

the ratio R becomes

$$R = \int\int_{G^*}\int\int \exp[E-F]\Big[\prod_{i=1}^{2} H_i(h_i)\Big] \prod_{i=1}^{2} \mu_1^{(i)}(dh_i)\mu_2^{(i)}(dk_i) \,. \tag{2.17}$$

To evaluate (2.17), we first evaluate

with (x,u), y, and v independent, $x{:}p_2 \times p_1$, $y{:}n_1 \times p_1$, $u{:}n \times p_2$ and $v{:}m \times p_2$. Then, the data in (2.7) corresponds to the data in (2.2) in terms of the new coordinates (x,y,u,v). Consider the group $\tilde{G} \equiv G\ell(p_1) \times G\ell(p_2) \times \mathfrak{O}(n_1) \times \mathfrak{O}(n) \times \mathfrak{O}(m)$ acting on (x,y,u,v) by

$$(2.8)\qquad (g_1,g_2,g_3,g_4,g_5)(x,y,u,v) = (g_2xg_1',\ g_3yg_1',\ g_4ug_2',\ g_5vg_2')$$

with $g_1 \in G\ell(p_1)$, $g_2 \in G\ell(p_2)$, $g_3 \in \mathfrak{O}(n_1)$, $g_4 \in \mathfrak{O}(n)$, and $g_5 \in \mathfrak{O}(m)$. Clearly, the group action of $\tilde{G}$ on (x,y,u,v) corresponds to the group action of G on $(S_{11.2},S_{21},S_{22},V)$ defined by (1.5).

Now, let $\tilde{\mathcal{D}}_\alpha$ be the class of all level α test functions [functions of (x,y,u,v)] for $H_0{:}\beta = 0$ versus $H_1{:}\beta \neq 0$ which are invariant under $\tilde{G}$. The mapping $S_{11.2} = y'y$, $S_{21} = x$, $S_{22} = u'u$, $V = v'v$ determines a 1-1 correspondence between test functions in $\mathcal{D}_\alpha$ and test functions in $\tilde{\mathcal{D}}_\alpha$. Thus, it is sufficient to verify (2.3) for test functions in $\tilde{\mathcal{D}}_\alpha$.

The joint density of (x,y,u,v), with respect to Lebesgue measure, is

$$(2.9)\qquad f(x,y,u,v|\beta) = c|u'u|^{-\frac{p_1}{2}} \exp[-\tfrac{1}{2}\mathrm{tr}[x'(u'u)^{-1}x + y'y]$$
$$+ \mathrm{tr}\beta'x - \tfrac{1}{2}\mathrm{tr}\beta'u'u\beta - \tfrac{1}{2}\mathrm{tr}(u'u + v'v)]$$

where c is a constant. Now, let P_β^T denote the probability measure of a maximal invariant, T, under the action of $\tilde{G}$ on (x,y,u,v). From a result due to Wijsman (1966), the Radon-Nikodym derivative $R = dP_\beta^T/dP_0^T$ can be written as

$$\begin{array}{lll} & S_{11.2} & \sim W(I_{p_1}, p_1, n_1) \\ (2.2) & \left\{ \begin{array}{l} S_{21} | S_{22} \\ \\ S_{22} \end{array} \right. & \begin{array}{l} \sim N(S_{22}\beta, S_{22} \times I_{p_1}) \\ \\ \sim W(I_{p_2}, p_2, n) \end{array} \\ & V & \sim W(I_{p_2}, p_2, m) \ . \end{array}$$

<u>Theorem 2.1</u>: For $\phi \in \mathcal{D}_\alpha$, the power function of ϕ, say $\pi(\phi, \beta) \equiv \pi(\phi; \delta_1, \ldots, \delta_t)$, is given by

$$(2.3) \qquad \pi(\phi; \delta_1, \ldots, \delta_t) = \alpha + \left(\sum_{i=1}^{t} \delta_i\right) B(\phi) + o\left(\sum_{1}^{t} \delta_i\right)$$

where

$$(2.4) \qquad B(\phi) = a_1 \mathcal{E}_0[\phi(a_0 \operatorname{tr} S_{11}^{-1} S_{12} (S_{22} + V)^{-1} S_{21} - \operatorname{tr} S_{22} (S_{22} + V)^{-1})]$$

with $a_0 = \frac{n}{p_1}$ and $a_1 = \frac{m+n}{2p_2}$. $\mathcal{E}_0$ denotes expectation when β in (2.2) is zero. The remainder term $o(\sum_{i=1}^{t} \delta_i)$ is uniform in ϕ, i.e.,

$$(2.5) \qquad \lim_{\sum \delta_i \to 0} \ \sup_{\phi \in \mathcal{D}_\alpha} \frac{o(\sum_{1}^{t} \delta_i)}{\sum_{1}^{t} \delta_i} = 0 \ .$$

The proof of Theorem 2.1 is rather lengthy and is deferred to the end of this section. An immediate consequence of Theorem 2.1 is

<u>Theorem 2.2</u>: Define the level α-test function $\phi^* \in \mathcal{D}_\alpha$ by

$$(2.6) \qquad \phi^* = \begin{cases} 1 & \text{if} \quad a_0 \operatorname{tr} S_{11}^{-1} S_{12} (S_{22} + V)^{-1} S_{21} - \operatorname{tr} S_{22} (S_{22} + V)^{-1} > k \\ 0 & \text{otherwise} \end{cases}$$

$$(1.19) \qquad R(u,w,\delta_0) \equiv \frac{g(u,w;\delta_0)}{g(u,w;0)} .$$

But, it is not hard to see that such tests depend on the particular value δ_0. Hence there can be no uniformly most powerful invariant test of H_0 versus H_1.

conditionally independent (given S_{22}), the joint conditional density is $S_{22}h_m(S_{22}w)\ q(u;S_{22}\delta)$. Multiplying this by the density of S_{22} (with $B_{22} = 1$) and integrating out S_{22} yields

(1.15) $g(u,w;\delta) =$

$$\sum_{j=0}^{\infty} \frac{\Gamma(\frac{m+n}{2} + j)}{\Gamma(\frac{m}{2})\ \Gamma(\frac{n}{2})}\ \frac{\delta^j}{j!}\ \frac{w^{\frac{m}{2} - 1}}{(1 + \delta + w)^{\frac{m+n}{2} + j}}\ g_{p-1+2j,n-p+1}(u); \quad u > 0 ,$$

as the joint density of U and W. Since the marginal density of W is $g_{m,n}(w)$, the conditional density of U given W is

(1.16) $g(u|w,\delta) =$

$$\sum_{j=0}^{\infty} \frac{\Gamma(\frac{m+n}{2} + j)}{\Gamma(\frac{m+n}{2})\ j!}\ \frac{(\frac{\delta}{1 + w})^j}{(1 + \frac{\delta}{1+w})^{\frac{m+n}{2} + j}}\ q_{p-1+2j,n-p+1}(u), \quad u > 0 .$$

<u>Proposition 1.2</u>: For each fixed value of w, $q(u|w,\delta)$ has a monotone likelihood ratio (MLR) in u and δ.

<u>Proof</u>: For w fixed, the function $(\frac{\delta}{1+w})^j/(1 + \frac{\delta}{1+w})^{\frac{m+n}{2} + j}$ has a MLR in δ and j. Also $q_{p-1+2j,n-p+1}(u)$ has a MLR in j and u. By a result due to Karlin (1956), $q(u|w,\delta)$ has a MLR in u and δ. This completes the proof.

Now let $C_\alpha(W)$ be the class of tests of H_0 versus H_1 which are invariant under G and which have conditional level α for each fixed value of W. Since all tests which are invariant under G can be written as functions of (U, W), we restrict our attention to the density of (U, W) when discussing the power of tests in $C_\alpha(W)$.

When $p_2 = 1$, then $|I - S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}| = 1 - R^2$ where $R^2 = S_{22}S_{11}^{-1}S_{12}/S_{22}$ is the multiple correlation coefficient.

For the remainder of this section we treat only the case $p_2 = 1$.

Proposition 1.1: Under the action of G defined by (1.5) a maximal invariant in the sample space is

$$(1.7) \qquad (U, W) \equiv \left(\frac{S_{21}S_{11.2}^{-1}S'_{21}}{S_{22}}, \frac{V}{S_{22}}\right)$$

and a maximal invariant in the parameter space is

$$(1.8) \qquad \delta \equiv B_{22}^{-1}B_{21}B_{11}^{-1}B'_{21} \geq 0 \; .$$

Proof: The proof of this proposition is routine and is omitted.

To study invariant tests of H_0 versus H_1, we now need to derive the joint distribution of U and W. The conditional distribution of (U, W) given S_{22} is first calculated and then S_{22} is integrated out. Since (U, W) is a maximal invariant under G, it suffices to find the distribution of (U, W) under the assumption that $B_{11} = I_{p-1}$ and $B_{22} = 1$ which we now assume. Conditional on S_{22}, $V/S_{22} \sim \frac{1}{S_{22}} \cdot \chi_m^2$ where "χ_m^2" denotes a chi-square random variable with m degrees of freedom. Since $B_{11} = I_{p-1}$ and $B_{22} = 1$,

$$(1.9) \qquad \frac{S_{21}}{\sqrt{S_{22}}}\Big| S_{22} \sim N(\sqrt{S_{22}}\,\Sigma_{21}, I_{p-1})$$

$$S_{11.2} \sim W(I_{p-1}, p-1, n-1) \; .$$

Thus, conditional on S_{22},

$$U = \frac{S_{21}}{\sqrt{S_{22}}} S_{11.2}^{-1} \frac{S'_{21}}{\sqrt{S_{22}}} \sim \frac{\chi_{p-1}^2\,(S_{22}\delta)}{\chi_{n-p+1}^2}$$

§ 1: Invariant Tests and the LRT for $p_2 = 1$.

In reduced form, the testing problem under consideration is the following. Given independent S and V,

$$(1.1)\qquad \begin{cases} S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \sim W(\Sigma, p, n) \\ \\ V \sim W(\Sigma_{22}, p, n) \end{cases}$$

we want to test $H_0 : \Sigma_{12} = 0$ versus $H_1 : \Sigma_{12} \neq 0$. Here, S_{11} is $p_1 \times p_1$, S_{22} is $p_2 \times p_2$ and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ is partitioned in the same way as S. Let G be the group of transformations defined by

$$(1.2)\qquad G = \left\{ A \;\middle|\; \begin{array}{l} A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix} : p \times p, \quad A_{11} : p_1 \times p_1 \\ A_{22} : p_2 \times p_2, \quad \det A_{ii} \neq 0, \quad i = 1,2 \end{array} \right\},$$

with the group action given by

$$(1.3)\qquad \begin{cases} S \to ASA' \\ V \to A_{22} V A_{22}' \\ \Sigma \to A \Sigma A' \\ \Sigma_{22} \to A_{22} \Sigma_{22} A_{22}' \end{cases} .$$

The testing problem is clearly invariant under G.

It is convenient to transform the observations S and V to other variables. Let $S_{11.2} = S_{11} - S_{12} S_{22}^{-1} S_{21}$ and $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$. Recall that (Eaton (1972))

Reducing the data first by sufficiency and then by translations (under which the testing problem is invariant), we have the statistics

$$(0.1)\qquad \begin{cases} S = \sum_1^N (X_i - \bar{X})(X_i - \bar{X})' \\ V = \sum_1^M (Y_i - \bar{Y})(Y_i - \bar{Y})' \end{cases}$$

where S and V are independent, $S \sim W(\Sigma, p, n)$ and $V \sim W(\Sigma_{22}, p_2, m)$ with $n = N - 1$ and $m = M - 1$. "$S \sim W(\Sigma, p, n)$" means that S has a p-dimensional Wishart distribution with n-degrees of freedom and expectation $n\Sigma$. Our problem is to test H_0: $\Sigma_{12} = 0$ versus H_1: $\Sigma_{12} \neq 0$. This problem has been considered previously by Lee and Geisser (1972) in the context of testing for "Rao simple structure" in the linear growth curve model. Lee and Geisser derived the likelihood ratio test and, surprisingly, the test is the same as if the additional data, V, was not present.

A slightly different testing problem which has the same structure as the above problem, after a reduction by sufficiency and translation, is the following. Consider a random sample $X_1, \ldots, X_N$ from a $N_p(\mu, \Sigma)$ distribution and partition Σ as above and let $\mu = \binom{\mu_1}{\mu_2}$ with $\mu_1 : p_1 \times 1$ and $\mu_2 : p_2 \times 1$. Also assume that $Y_1, \ldots, Y_M$, is a random sample (independent of the X's) from a $N_{p_2}(\mu_2, \Sigma_{22})$ distribution. Again it is desired to test H_0: $\Sigma_{12} = 0$ versus H_1: $\Sigma_{12} \neq 0$. Reducing the data by sufficiency we have $\bar{X} \sim N(\mu, \frac{1}{N}\Sigma)$,

$$S = \sum_1^n (X_i - \bar{X})(X_i - \bar{X})' \sim W(\Sigma, p, N-1) ,$$

$\bar{Y} \sim N(\mu_2, \frac{1}{m}\Sigma_{22})$ and $\tilde{V} = \sum_1^m (Y_i - \bar{Y})(Y_i - \bar{Y})' \sim W(\Sigma_{22}, p_2, m-1)$. The problem

References

[1] Eaton, M. L. (1972). Multivariate Statistical Analysis. Institute of Math. Statist., Univ. of Copenhagen, Copenhagen, Denmark.

[2] Giri, N. and Kiefer, J. (1964). Local and asymptotic minimax properties of a normal multivariate testing problem. Ann. Math. Statist., 35, 21-35.

[3] James, A. T. (1961). The distributions of noncentral means with known covariance. Ann. Math. Statist., 32, 874-882.

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

[5] Kariya, T. (1974). Hypothesis testing problems in the general MANOVA model. Ph.D. Thesis. Univ. of Minnesota.

[6] Karlin, S. (1956). Decision theory for Polya type distributions. Case of two actions, I. Third Berk. Symp. Math. Statist. Prob. 1, 115-129, Univ. of California.

[7] Lee, J. C. and Geisser, S. (1972). Growth curve prediction. Sankhyā, A34, 393-412.

[8] Lehmann, E. L. (1959). Testing Statistical Hypotheses. Wiley, New York.

[9] Schwartz, R. (1967). Local minimax tests. Ann. Math. Statist., 38, 340-360.

[10] Wijsman, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. Fifth Berk. Symp. Math. Statist. Prob. 1, 389-400, Univ. of California.

Thus, for $\operatorname{tr} \beta\beta' < \frac{1}{4p}$,

$$(2.27) \qquad \begin{aligned} |IV| \le |V| + |VI| + |VII| &\le [2p \operatorname{tr} \beta\beta']^2 \exp[\tfrac{1}{4} \operatorname{tr} h_2h_2'] \\ &+ [4p \operatorname{tr} \beta\beta']^2 \exp[\tfrac{1}{4} \operatorname{tr} h_1h_1' + \tfrac{1}{4} \operatorname{tr} h_2h_2'] \\ &+ [4p \operatorname{tr} \beta\beta']^2 \exp[\tfrac{1}{4} \operatorname{tr} h_1h_1' + \tfrac{3}{8} \operatorname{tr} h_2h_2'] \\ &\equiv [\operatorname{tr} \beta\beta']^2 K(h_1,h_2) \ . \end{aligned}$$

Integrating (2.21) and (2.27) with respect to $\mu_2^{(1)}(dk_1)$ yields

$$(2.28) \qquad \begin{cases} \int II \ \mu_2^{(1)}(dk_1) = \dfrac{1}{2p_2} \operatorname{tr} \bar{g}_2u'u\bar{g}_2' \operatorname{tr} h_2h_2' \beta\beta' \\ \int III \ \mu_2^{(1)}(dk_1) = \dfrac{1}{2p_1p_2} \operatorname{tr} \bar{g}_1x'\bar{g}_2'x\bar{g}_1' \operatorname{tr} h_1'\beta'h_2h_2'\beta h_1 \\ \int |IV| \ \mu_2^{(1)}(dk_1) \le [\operatorname{tr} \beta\beta']^2 K(h_1,h_2) \quad \text{for} \quad \operatorname{tr} \beta\beta' < \dfrac{1}{4p} \ . \end{cases}$$

Now, integrating the three members of (2.28) with respect to $\prod_{i=1}^{2} H_i(h_i)\mu_1^{(i)}(dh_i)$, and thereby completing the integration in (2.17), we have, for $\operatorname{tr} \beta\beta'$ small,

$$(2.29) \qquad \begin{aligned} R = 1 &+ \frac{n(n+m)}{2p_1p_2} \operatorname{tr} \bar{g}_1x'\bar{g}_2'\bar{g}_2x\bar{g}_1' \operatorname{tr} \beta\beta' \\ &- \frac{(n+m)}{2p_2} \operatorname{tr} \bar{g}_2u'u\bar{g}_2' \operatorname{tr} \beta\beta' + o(\operatorname{tr} \beta\beta') \end{aligned}$$

where the remainder term is uniform in (x,y,u,v).

Since $\phi \in \tilde{\mathfrak{S}}_{\alpha}$ is invariant, it is a function of any maximal invariant. Thus, for $\sum_1^t \delta_i = \operatorname{tr}\beta\beta'$ small, we have

$$(2.18) \qquad A \equiv \int_{\mathcal{O}(p_2)} \exp[E-F]\, \mu_2^{(2)}(dk_2) = \int_{\mathcal{O}(p_2)} \sum_{j=0}^{\infty} \frac{(E-F)^j}{j!}\, \mu_2^{(2)}(dk_2)\ .$$

Since the terms in odd powers of E vanish, we have

$$(2.19) \quad \begin{aligned} A &= \int_{\mathcal{O}(p_2)} [1 - F + \tfrac{1}{2}E^2 + \tfrac{1}{2}F^2 + \sum_{j=3}^{\infty} \frac{(E-F)^j}{j!}]\mu_2^{(2)}(dk_2) \\ &= 1 - \int_{\mathcal{O}(p_2)} F\, \mu_2^{(2)}(dk_2) + \tfrac{1}{2}\int_{\mathcal{O}(p_2)} E^2\, \mu_2^{(2)}(dk_2) \\ &\qquad + \int_{\mathcal{O}(p_2)} [\tfrac{1}{2} F^2 + \sum_{j=3}^{\infty} \frac{(E-F)^j}{j!}]\, \mu_2^{(2)}(dk_2) \\ &\equiv 1 - II + III + IV\ . \end{aligned}$$

For any $q \times q$ matrices B_1 and B_2, recall that

$$(2.20) \quad \begin{cases} \int_{\mathcal{O}(q)} \operatorname{tr} B_1 \Gamma B_2 \Gamma'\ \lambda(d\Gamma) = \frac{1}{q} \operatorname{tr} B_1 \operatorname{tr} B_2 \\ \int_{\mathcal{O}(q)} (\operatorname{tr} B_1 \Gamma)^2 \lambda(d\Gamma) = \frac{1}{q} \operatorname{tr} B_1 B_1' \end{cases}$$

where λ is invariant probability measure on $\mathcal{O}(q)$ (see James (1961, 1964)). Using (2.20), we have

$$(2.21) \quad \begin{cases} II = \frac{1}{2p_2} \operatorname{tr} \bar{g}_2 u' u \bar{g}_2' \operatorname{tr} h_2 h_2'\ \beta\beta' \\ III = \frac{1}{2p_2} \operatorname{tr} \bar{g}_1 x' \bar{g}_2' \bar{g}_2\ x \bar{g}_1'\ k_1' h_1'\ \beta'\ h_2 h_2'\ \beta\ h_i k_i\ . \end{cases}$$

Further, a bit of manipulation shows that

$$(2.10) \qquad R = \frac{\int_{\tilde{G}} f(g(x,y,u,v)|\beta)|J|^{-1}\nu(dg)}{\int_{\tilde{G}} f(g(x,y,u,v)|0)|J|^{-1}\nu(dg)}$$

where ν is a left invariant measure on $\tilde{G}$ and $|J|$ is the Jacobian of the transformation defined by (2.8). We note here that the reason for transforming to the new variables (x,y,u,v) is so that the conditions necessary to use the representation (2.10) could be easily verified.

The remainder of the proof is primarily concerned with the evaluation of (2.10) for $\Sigma\delta_i$ small. For $g = (g_1,g_2,g_3,g_4,g_5) \in \tilde{G}$ as in (2.8), the Jacobian of the transformation (2.8) is

$$(2.11) \qquad |J| = |g_1'g_1|^{-\frac{n}{2}}|g_2'g_2|^{-\frac{p_1+n+m}{2}} .$$

Now, select $\bar{g}_1 \in G_T^+(p_1)$ such that $\bar{g}_1(x'(u'u)^{-1}x + y'y)\bar{g}_1' = I_{p_1}$ and select $\bar{g}_2 \in G_T^+(p_2)$ such that $\bar{g}_2(u'u + v'v)\bar{g}_2' = I_{p_2}$. Substituting $(g_1\bar{g}_1, g_2\bar{g}_2)$ for (g_1,g_2) without changing the value of R, the numerator in (2.10) becomes, after some cancellation,

$$(2.12) \qquad N = \int\int_{G\ell(p_1)\times G\ell(p_2)} |J_1|\ |J_2|\exp[-\tfrac{1}{2}\mathrm{tr}\ g_1'g_1 - \tfrac{1}{2}\mathrm{tr}\ g_2'g_2]$$
$$\exp[\mathrm{tr}\ \beta'g_2\bar{g}_2x\bar{g}_1'g_1' - \tfrac{1}{2}\mathrm{tr}\ \beta'g_2\bar{g}_2(u'u)\bar{g}_2'g_2\beta]\nu_1(dg_1)\nu_2(dg_2) ,$$

where $|J_1| = |g_1'g_1|^{\frac{n(1)}{2}}$, $|J_2| = |g_2'g_2|^{\frac{n(2)}{2}}$, ν_i is left invariant measure on $G\ell(p_i)$, $i = 1,2$, and $n(1) = n$, $n(2) = m+n$. For $g_i \in G\ell(p_i)$, where k is chosen to make ϕ^* level α. Then ϕ^* is the unique locally best invariant test of H_0. Further, ϕ^* is admissible in the class $\mathcal{D}_\alpha$.

Proof: The assertions follow from the generalized Neyman-Pearson Lemma given in Lehmann (1959).

Since the test ϕ^* is locally most powerful in $\mathcal{D}_\alpha$, it is natural to ask about the local minimaxity of ϕ^* within the class of all level α tests. (See Giri-Kiefer (1964) for the definition and sufficient conditions under which a given test is locally minimax.) When $p_1 = p_2 = 1$, it is clear that ϕ^* is locally minimax. However, we have been unable to show that ϕ^* is locally minimax when either p_1 or p_2 is bigger than 1.

Proof of Theorem 2.1:

Throughout this proof, $G_T(r)$ will denote the group of $r \times r$ lower triangular matrices with positive diagonal elements, $\mathcal{O}(r)$ will denote the group of $r \times r$ orthogonal matrices, and $G\ell(r)$ will denote the group of $r \times r$ non-singular matrices. Recall that if $\eta \sim W(\Sigma, r, k)$ then $\eta \sim \xi'\xi$ where ξ is $k \times r$ and $\xi \sim N(0, I_k \times \Sigma)$. Using this decomposition of the Wishart distribution, let $S_{11.2} = y'y$, $S_{21} = x$, $S_{22} = u'u$ and $V = v'v$ where the joint distribution of (x,y,u,v) is given by

$$(2.7) \quad \left\{ \begin{array}{ll} x|u & \sim N(u'u\beta, (u'u) \times I_{p_1}) \\ y & \sim N(0, I_{n_1} \times I_{p_1}) \\ u & \sim N(0, I_n \times I_{p_2}) \\ v & \sim N(0, I_m \times I_{p_2}) \end{array} \right.$$

§ 2: A Locally Most Powerful Invariant Test.

In the notation of the previous section, the data for the testing problem under consideration is $(S_{11.2}, S_{21}, S_{22}, V)$ where

$$
(2.1)\quad
\begin{array}{ll}
S_{11.2} & \sim W(\Sigma_{11.2}, p_1, n_1); \quad n_1 = n - p_2 \\
\left\{\begin{array}{l} S_{21}|S_{22} \\ S_{22} \end{array}\right. & \begin{array}{l} \sim N(S_{22}\Sigma_{22}^{-1}\Sigma_{21},\ S_{22} \times \Sigma_{11.2}) \\ \sim W(\Sigma_{22}, p_2, n) \end{array} \\
V & \sim W(\Sigma_{22}, p_2, m)
\end{array}
$$

and $S_{11.2}$, (S_{21}, S_{22}) and V are mutually independent. The problem is to test H_0: $\Sigma_{21} = 0$ versus H_1: $\Sigma_{21} \neq 0$. This problem is invariant under the group G defined by (1.2) with the group action being defined by (1.3) (or in terms of the new variables by (1.5)). Since p_1 and p_2 are now arbitrary, an analytically tractable maximal invariant seems difficult to obtain. A maximal invariant parameter is the set of ordered characteristic roots of $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ -- say, δ_i, $i = 1,\ldots,t$ where $t = \min\{p_1, p_2\}$, with $\delta_1 \geq \delta_2 \geq \ldots \geq \delta_t \geq 0$.

Let $\mathcal{D}_\alpha$ be the set of all G-invariant level α test functions for H_0 versus H_1. Only test functions in $\mathcal{D}_\alpha$ will be considered in this section. If $\phi \in \mathcal{D}_\alpha$, it is well known (Lehmann (1959)) that the power function of ϕ will be a function of a maximal invariant parameter. Let $\beta: p_2 \times p_1$ be such that $\beta_{ii} = \sqrt{\delta_i}$ for $i = 1,\ldots,t$ with the remaining elements of β being zero. Since our concern is the computation of power functions for invariant tests $\phi \in \mathcal{D}_\alpha$, we can, without loss of generality, assume that the data is

Theorem 1: Let $\phi^*(U)$ be the test defined by

$$(1.17) \qquad \phi^*(U) = \begin{cases} 1 & \text{if } U > c \\ 0 & \text{if } U \le c \end{cases}$$

where c is chosen to make ϕ^* level α. Then ϕ^* is uniformly most powerful in the class $C_\alpha(W)$.

Proof: Let $\varphi(U,W) \in C_\alpha(W)$. Then $\mathcal{E}_{\delta=0}(\varphi(U,W)|W) = \alpha$ a.e. (W). Since the conditional density of U given W has a MLR, we have that

$$(1.18) \qquad \mathcal{E}_\delta(\phi^*(U)|W) \ge \mathcal{E}_\delta(\phi(U,W)|W) \ .$$

Integrating both sides of (1.18) with respect to the distribution of W yields the result.

A simple calculation shows that $U = R^2/(1-R^2)$ where $R^2 = S_{21}S_{11}^{-1}S_{12}/S_{22}$ so the test ϕ^* is equivalent to the likelihood ratio test. Since W is an ancillary statistic, one might argue, using the Principle of Conditionality, that after a reduction by invariance, testing should be done conditionally on W.

In the next section, we show that the test ϕ^* can be dominated locally (in terms of power). More specifically, a locally most powerful invariant test is constructed for general p_1 and p_2. As will be seen, this test depends on both U and W and so is not equivalent to ϕ^*.

Let C_α be the class of all level α invariant tests. To show that a uniformly most powerful test in C_α does not exist we argue as follows. Fix $\delta_0 > 0$. For testing $H_0:\delta = 0$ versus $H_1:\delta = \delta_0$, the most powerful invariant test rejects for large values of

where χ^2_{n-p+1} and $\chi^2_{p-1}(S_{22}\delta)$ are independent and $\chi^2_{p-1}(S_{22}\delta)$ denotes a non-central chi-square random variable with non-centrality parameter $S_{22}\delta$.

The density of $\chi^2_k(\tau)$ is given by

$$(1.10) \qquad f(x;\tau) = \sum_{j=0}^{\infty} \frac{e^{-\frac{\tau}{2}}}{j!} \left(\frac{\tau}{2}\right)^j h_{k+2j}(x)$$

where

$$(1.11) \qquad h_\alpha(x) = \frac{x^{\frac{\alpha}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{\alpha}{2}}\,\Gamma(\alpha/2)}\ , \qquad x > 0$$

is the density of a χ^2_α random variable. Also, the density of $F_{\alpha,\beta} \equiv \frac{\chi^2_\alpha}{\chi^2_\beta}$, where χ^2_α and χ^2_β are independent, is

$$(1.12) \qquad g_{\alpha,\beta}(z) = \frac{\Gamma(\frac{\alpha+\beta}{2})}{\Gamma(\alpha/2)\Gamma(\beta/2)} \, \frac{y^{\alpha/2-1}}{(1+y)^{\frac{\alpha+\beta}{2}}}\ , \qquad y > 0\ .$$

From (1.10) and (1.12) it follows that the density of $\chi^2_\alpha(\tau)/\chi^2_\beta$ is

$$(1.13) \qquad g(z;\tau) = \sum_{j=0}^{\infty} \frac{e^{-\frac{\tau}{2}}}{j!} \left(\frac{\tau}{2}\right)^j \; g_{\alpha+2j,\beta}(z); \qquad z > 0\ .$$

Thus, the conditional density of U, given S_{22}, is

$$(1.14) \qquad q(u;S_{22}\delta) = \sum_{j=0}^{\infty} \frac{e^{-\frac{S_{22}\delta}{2}}}{j!} \left(\frac{S_{22}\delta}{2}\right)^j \; g_{p-1+2j,n-p+1}(u); \qquad u > 0\ .$$

Further, the conditional density of $W \equiv V/S_{22}$, given S_{22} is $S_{22}h_m(S_{22}w)$ where h_m is given by (1.11). Since U and W are

$$(1.4)\quad \begin{cases} S_{11.2} \sim W(\Sigma_{11.2}, p_1, n-p_2) \\ S_{21}|S_{22} \sim N(S_{22}\Sigma_{22}^{-1}\Sigma_{21},\ S_{22}\times\Sigma_{11.2}) \\ S_{22} \sim W(\Sigma_{22}, p_2, n)\ . \end{cases}$$

Also, $S_{11.2}$ is independent of the pair (S_{21}, S_{22}). The second line in (1.4) means that the conditional distribution of S_{21} given S_{22} is multivariate normal with mean matrix $S_{22}\Sigma_{22}^{-1}\Sigma_{21}$ and covariance operator $S_{22}\times\Sigma_{11.2}$ - the Kronecker product of S_{22} and $\Sigma_{11.2}$. Let $B_{11}=\Sigma_{11.2}$, $B_{21}=\Sigma_{22}^{-1}\Sigma_{21}$ and $B_{22}=\Sigma_{22}$, so our testing problem is now $H_0: B_{21}=0$ versus $H_1: B_{21}\neq 0$. The group action on the new variables induced by (1.3) is

$$(1.5)\quad \begin{aligned} S_{11.2} &\to A_{11}S_{11.2}A_{11}' \\ S_{21} &\to A_{22}S_{21}A_{11}' \\ S_{22} &\to A_{22}S_{22}A_{22}' \\ V &\to A_{22}VA_{22}' \\ B_{11} &\to A_{11}B_{11}A_{11}' \\ B_{21} &\to (A_{22}')^{-1}B_{21}A_{11}' \\ B_{22} &\to A_{22}B_{22}A_{22}'\ . \end{aligned}$$

With the observations given by (1.4) plus V, it is easy to calculate the maximum likelihood estimators for the parameters B_{11}, B_{21} and B_{22} under H_0 and H_1, and thus compute the likelihood ratio test. The likelihood ratio test rejects H_0 for small values of the statistic

$$(1.6)\quad |I - S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}|\ .$$

is invariant under $\bar{X} \to \bar{X} + a$, $S \to S$, $\bar{Y} \to \bar{Y} + a_2$, $\tilde{V} \to \tilde{V}$ where $a = \binom{a_1}{a_2}$, $a_1 = p_1 \times 1$, $a_2 : p_2 \times 1$. A maximal invariant under this group of translations is $(S, V, \bar{Y} - \bar{X}_{(2)})$ where $\bar{X} = \binom{\bar{X}_{(1)}}{\bar{X}_{(2)}}$ with $\bar{X}_{(1)} : p_1 \times 1$ and $\bar{X}_{(2)} : p_2 \times 1$. But

$$\bar{Y} - \bar{X}_{(2)} \sim N_{p_2}(0, (\tfrac{1}{N} + \tfrac{1}{m}) \Sigma_{22}) \quad \text{so}$$

$$Z = (\tfrac{1}{N} + \tfrac{1}{m})^{-\frac{1}{2}}(\bar{Y} - \bar{X}_{(2)}) \sim N(0, \Sigma_{22}) \ .$$

Reducing the data $(S, V, \bar{Y} - \bar{X}_{(2)})$ by sufficiency, we have the sufficient statistic $(S, \tilde{V} + ZZ') \equiv (S, V)$. Here $S \sim W(\Sigma, p, N-1)$ and is independent of $V \sim W(\Sigma_{22}, p_2, m+1)$. In this reduced form, the problem for testing $H_0 : \Sigma_{12} = 0$ versus $H_1 : \Sigma_{12} \neq 0$ is identical to the problem when the extra observations have a mean which is unrelated to the mean of the original sample.

In Section 1, we show that for $p_2 = 1$ the likelihood ratio test is the uniformly most powerful invariant conditional level-α test. The conditioning is with respect to a natural ancillary statistic. In addition, it is argued that a uniformly most powerful invariant (under the appropriate group of linear transformations leaving the problem invariant) test does not exist.

A locally most powerful invariant test for H_0 versus H_1 is derived in Section 2. This test is seen to depend on an ancillary statistic and the test is different from the likelihood ratio test. A representation theorem for the probability ratio of the density of a maximal invariant due to Wijsman (1967) is used to derive the locally most powerful invariant test.

§ 0: Introduction and Notation.

Let $X_1,\ldots,X_N$ be a random sample from a p-dimensional multivariate normal distribution $N_p(\mu, \Sigma)$ and partition Σ as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}; \quad \Sigma_{11}: p_1 \times p_1, \quad \Sigma_{22}: p_2 \times p_2$$

with $p = p_1 + p_2$. For testing $H_0: \Sigma_{12} = 0$ versus $H_1: \Sigma_{12} \neq 0$, when $p_2 = 1$, it is well known that the best invariant test (which is the likelihood ratio test) rejects H_0 for large values of the sample multiple correlation coefficient between the first $p - 1$ coordinates and the last coordinate. For general p_1 and p_2, a variety of invariant tests have been proposed -- all of which are functions of the sample eigenvalues of $S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ where S is the sample covariance matrix based on $X_1,\ldots,X_n$ and

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \quad S_{11}:p_1 \times p_1, \quad S_{22}:p_2 \times p_2 .$$

In particular, Schwartz (1967) has shown that the test which rejects H_0 for large values of $\mathrm{tr}S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}$ (the Pillai trace test) is locally most powerful invariant and locally minimax in some special cases.

The purpose of the present paper is to consider the above testing problem when one has additional observations on the last p_2 coordinates. In addition to $X_1,\ldots,X_N$ assume that $Y_1,\ldots,Y_M$ is a random sample (independent of the X's) from an p_2-dimensional multivariate normal distribution - $N_{p_2}(\tau, \Sigma_{22})$. Presumably, the Y's can be used to help estimate Σ_{22} and thus be used to construct a more powerful test of H_0.

ABSTRACT

Suppose $S : p \times p$ has a Wishart distribution with n degrees of freedom and expectation $n\Sigma$, $n \geq p$. Write

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \text{ where}$$

$\Sigma_{11}: p_1 \times p_1$ and $\Sigma_{22}: p_2 \times p_2$, $p_1 + p_2 = p$. Also suppose $V : p_2 \times p_2$ has a Wishart distribution with m-degrees of freedom and expectation $m\Sigma_{22}$. For testing $H_0 : \Sigma_{12} = 0$ versus $H_1 : \Sigma_{12} \neq 0$, we derive a locally most powerful invariant test and show that this test is different from the likelihood ratio test.