The Annals of Statistics
1973, Vol. 1, No. 3, 682–699

APPLICATION OF THE THEORY OF PRODUCTS OF PROBLEMS TO CERTAIN PATTERNED COVARIANCE MATRICES[1]

By Steven F. Arnold

Lawrence University

This paper describes a general method for deriving optimal procedures for problems where the covariance matrices are patterned under both null and alternative hypotheses. The pattern considered in this paper was first suggested by Olkin (1970) and is a generalization of the intraclass correlation model of Wilks (1946) and arises in the study of interchangeable random variables. We prove a theorem showing how we can transform most such problems to products of problems where the covariance matrices are unpatterned. This theorem is applied to two problems, the multivariate analysis of variance problem and the multivariate classification problem where in both cases the covariance matrix is assumed patterned. We use theorems about products to derive optimal procedures for these problems. We then look at Olkin's pattern for the mean vector, and show that most problems where both the mean vector and covariance matrix are patterned can be transformed to a product of problems, one of which is trivial. The same two examples are studied where now both mean vectors and covariance matrices are assumed patterned. We also consider the problem of testing that the mean vector is patterned when we know the covariance matrix is.

1. Introduction.

1.1. This paper is concerned with normal testing problems involving interchangeable random variables. $X_1, \cdots, X_k$ are said to be *interchangeable* if for any permutation σ of $1, \cdots, k$, the distribution of $X_1, \cdots, X_k$ is the same as the distribution of $X_{\sigma(1)}, \cdots, X_{\sigma(k)}$. Interchangeable random variables arise in many situations of dependent sampling such as the problem where the pollster sends cards to all the people on one street. His observations would not be independent but it might be reasonable to assume that they are interchangeable. Another rather different example is the following. A biologist who is studying kidneys uses measurements on both kidneys of each rat instead of only using one kidney. The two kidneys would probably not be independent, but might well be interchangeable.

Let $X' = (X_1', \cdots, X_k')$ where X_i is $p \times 1$. If X has a kp-variate normal distribution, then a necessary and sufficient condition that $X_1, \cdots, X_k$ be interchangeable is

$$(1.1) \qquad \mu = EX = \begin{pmatrix} \theta \\ \vdots \\ \theta \end{pmatrix}, \qquad \Sigma = E(X-\mu)(X-\mu)' = \begin{pmatrix} \Sigma_1 & \Sigma_2 & \cdots & \Sigma_2 \\ \Sigma_2 & \Sigma_1 & \cdots & \Sigma_2 \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_2 & \Sigma_2 & \cdots & \Sigma_1 \end{pmatrix},$$

Received May 1972; revised October 1972.

[1] Prepared under the auspices of National Science Foundation Grant GP 17172.

AMS 1970 *subject classifications.* 62H15.

Key words and phrases. Intraclass correlation, multivariate analysis of variance, classification.

for some $p \times 1$ vector θ and $p \times p$ symmetric materices Σ_1 and Σ_2. If μ and Σ satisfy (1.1) for some Σ_1, Σ_2, symmetric, and θ, we say μ has pattern B_k and Σ has pattern A_k. Hence, $X_1, \cdots, X_k$ are interchangeable if and only if μ has pattern B_k, Σ has pattern A_k.

These patterns are generalizations of the intraclass correlation model introduced in Wilks (1946) and extended in Olkin (1970). However, in this paper we consider somewhat different problems involving these patterns, in that we assume that the covariance matrix (and sometimes the means too) are patterned under both the null and alternative hypotheses. In the univariate case ($p = 1$), problems of this type have been considered in Geisser (1963), Srivastava (1965) and Krishnaiah and Pathak (1967).

Our approach to these problems is quite different from the approaches found in those papers. In Section 3 we prove the basic theorems of the paper which allow us to transform problems involving patterned covariance matrices to "products" of problems where the covariances are not assumed patterned. In Section 2 we state two theorems telling how to derive optimal procedures for a product from optimal procedures for the components. Then in Sections 4–6 we give three examples of problems that transform in this way: (i) testing that the mean has pattern B_k when the covariance matrix has pattern A_k (one of the problems first considered in the univariate case in Wilks (1946) and later in the multivariate case in Olkin (1970)); (ii) the multivariate analysis of variance problem when the covariance matrix is patterned (a generalization of the Hotelling's T^2 problem considered in the case $p = 1$ by Geisser (1963)); and, (iii) the multivariate classification problem when the covariance matrix is patterned. In Section 7 we prove a general theorem showing that problems where both the means and covariance matrices are patterned (i.e., involving interchangeable random variables) can be transformed to a product of a "trivial" problem and a problem identical to the original problem except that nothing is assumed patterned. In Section 2 we show that if we have a product of a non-trivial problem and a trivial one, then any optimal procedure for the non-trivial problem is optimal for the product. Therefore, Section 7 shows that if we have a problem where both means and covariance matrix are patterned then we can, in some sense, reduce it to an identical problem, except that nothing is patterned. We give two examples of this structure; reducing the multivariate analysis of variance and classification problems where both the means and covariance matrices are patterned to similar problems where they are not patterned. In Section 8, we indicate other problems that can be analyzed using the methods of this paper.

1.2. The following distributions are used in this paper. If μ is a $p \times r$ matrix and Σ is a $p \times p$ matrix, we write $X(p \times r) \sim N(\mu, \Sigma)$ to mean X is a $p \times r$ dimensional matrix whose columns are independently normally distributed with common covariance matrix Σ and $EX = \mu$. Therefore

$$\mathscr{L}(X) = (2\pi)^{-rp/2} |\Sigma|^{-r/2} \operatorname{etr} -\tfrac{1}{2}[\Sigma^{-1}(X - \mu)(X - \mu)'] ,$$

where $\operatorname{etr} A = \exp(\operatorname{tr} A)$. If $X(p \times r) \sim N(0, \Sigma)$, we say $W = XX'$ has a Wishart distribution with r degrees of freedom and write $W \sim W(r, \Sigma)$. If $X(1 \times n) \sim N(\mu, 1)$, we say $S = XX'$ has a noncentral χ^2 distribution with n degrees of freedom and noncentrality parameter $\delta = \mu\mu'$ and write $S \sim \chi_n^2(\delta)$. If $S_1 \sim \chi_n^2(\delta)$, $S_2 \sim \chi_m^2(0)$ we say $T = mS_1/nS_2$ has a noncentral F distribution with n and m degrees of freedom and noncentrality parameter δ and write $T \sim F(\delta; m, n)$. If $T \sim F(\delta; m, n)$ we say $U = mT(n + mT)^{-1}$ has a noncentral Beta distribution with $r = m/2$, $s = n/2$ degrees of freedom and noncentrality parameter $\lambda = \delta/2$ and write $U \sim \mathrm{Be}(\lambda; r, s)$.

$$\mathscr{L}(U) = e^{-\lambda} \sum_{t=0}^{\infty} \frac{\lambda^t}{t!} \frac{U^{r+t}(1-U)^s}{B(r+t, s)},$$

where $B(a, b)$ is the beta function.

2. Products of problems.

2.1. Before stating theorems about products we give a shorthand method of describing a general testing problem, and use it to define a product of problems. A testing problem P consists of the following three elements: an observed random variable X having density from a general class $D(\theta)$ (for example $N(\mu, \Sigma)$); a null set Ω; and an alternative set Θ. We use the following shorthand for P

$$\begin{aligned} P&: X \sim D(\theta), \\ H&: \theta \in \Omega, \\ A&: \theta \in \Theta. \end{aligned}$$

We make one convention for this notation. *All random variables are independent unless otherwise specified.* For example, we write the Hotelling's T^2 problem (testing $\mu = 0$ when Σ is unknown) in the following manner:

$$\begin{aligned} P&: X \sim N(\mu, \Sigma), \qquad W \sim W(n, \Sigma), \\ H&: \mu = 0, \qquad \Sigma > 0, \\ A&: -\infty < \mu < \infty, \qquad \Sigma > 0. \end{aligned}$$

(Throughout this paper we write $-\infty < \mu < \infty$ to mean that μ is unrestricted.)

So now we define a product of problems. Let P_1 and P_2 be the problems

$$\begin{aligned} P_1&: X_1 \sim D_1(\theta_1), & P_2&: X_2 \sim D_2(\theta_2), \\ H_1&: \theta_1 \in \Omega_1, & H_2&: \theta_2 \in \Omega_2, \\ A_1&: \theta_1 \in \Theta_1, & A_2&: \theta_2 \in \Theta_2. \end{aligned}$$

Then the *product* P of P_1 and P_2 (written $P = P_1 \times P_2$) is the problem

$$\begin{aligned} P&: X_1 \sim D_1(\theta_1), \qquad X_2 \sim D_2(\theta) \\ H_0&: \theta_1 \in \Omega_1, \qquad \theta_2 \in \Omega_2, \\ A_0&: \theta_1 \in \Theta_1, \qquad \theta_2 \in \Theta_2. \end{aligned}$$

That is, the product P is just the problem of testing P_1 and P_2 simultaneously and independently. We make the following three definitions for testing problems. A problem is *inclusive* if the null set, Ω, is a subset of the alternative set, Θ, and *symmetric* if Ω and Θ are mutually exclusive. For example the Hotelling's T^2 problem is inclusive, while the classification problem is symmetric. A problem is *simple* if the null set consists of only one point. If P is invariant under the group G, we write P/G for the problem P reduced by the group G. The following theorem gives a collection of straightforward results about products. Since much of it may be known and its proof is easy we omit it here. For details see Arnold (1970).

THEOREM A. *Let P_1 and P_2 be the problems given above and let $P = P_1 \times P_2$.*

(i) *If S_1 and S_2 are sufficient statistics for P_1 and P_2 respectively, then (S_1, S_2) is a sufficient statistic for P.*

(ii) *If Λ_1 and Λ_2 are the likelihood ratio test* (LRT) *functions for P_1 and P_2 respectively, then $\Lambda_1 \Lambda_2$ is the* LRT *function for P.*

(iii) *If f_1 and f_2 are the Bayes test functions for P_1 and P_2, with respect to the priors $Q_1(\theta_1)$ aud $Q_2(\theta_2)$, then $f = f_1 f_2$ is the Bayes test function with respect to the prior $Q(\theta_1, \theta_2) = Q_1(\theta_1) \times Q_2(\theta_2)$.*

(iv) *Let P_1 and P_2 be simple or symmetric problems and let $f_1(X_1)$ and $f_2(X_2)$ be unbiased test functions for P_1 and P_2. If $g(z_1, z_2)$ is an increasing function of z_1 and z_2, the $g(f_1(X_1), f_2(X_2))$ is an unbiased test function for P.*

(v) *Let $\Theta_1 \cup \Omega_1$ and $\Theta_2 \cup \Omega_2$ be partially ordered sets and let $f_1(X_1)$ and $f_2(X_2)$ have monotone power for these orderings for P_1 and P_2. If $g(z_1, z_2)$ is an increasing function of z_1 and z_2, then $g(f_1(X_1), f_2(X_2))$ has monotone power for the induced ordering $(\theta_{11}, \theta_{12}) \leqq (\theta_{21}, \theta_{22})$ if and only if $\theta_{11} \leqq \theta_{21}$ and $\theta_{12} \leqq \theta_{22}$.*

(vi) *If P_1 is invariant under G_1 and P_2 is invariant under G_2, then P is invariant under $G = G_1 \times G_2$ operating by $g(X_1, X_2) = (g_1(X_1), g_2(X_2))$ where $g = (g_1, g_2)$.*

(vii) *$P/G = P_1/G_1 \times P_2/G_2$. That is, P reduced by invariance is product of P_1 and P_2 reduced by invariance.*

One comment about A(ii) and A(iii). For this paper, the LRT function Λ and the Bayes test function $f(X)$ with respect to the prior $Q(\theta)$, are defined to be

$$\Lambda(X) = \frac{\sup_{\theta \in \Omega} p(X; \theta)}{\sup_{\theta \in \Theta} p(X; \theta)} , \quad f(X) = \frac{\int_\Omega p(X; \theta) Q(\theta)\, d\theta}{\int_\Theta p(X; \theta) Q(\theta)\, d\theta} . \tag{2.1}$$

Often, statisticians work with functions that are statistically equivalent to the LRT function or Bayes test functions (by statistically equivalent functions, we mean that one function is an increasing function of the other). For Theorems A(ii) and A(iii) we must use the functions defined in (2.1). If f_1 is statistically equivalent to f_2, and g_1 is statistically equivalent to g_2, there is no reason why $f_1 g_1$ would be statistically equivalent to $f_2 g_2$.

Three of the problems that we consider transform into a product where one problem is *trivial*, that is, the null set and the alternative set are the same. In this case we get more powerful results. (The proof of this theorem is again straightforward. Details can be found in Arnold (1970).)

THEOREM B. *Let $P = P_1 \times P_2$ where P_2 is trivial.*

(i) *If P_2 is simple then X_1 is a sufficient statistic for P.*

(ii) *If $\varphi(X_1)$ has any of the following ten properties for P_1, it has that property for P.*

P1. LRT,
P2. *Bayes*,
P3. *admissible*,
P4. *unbiased*,
P5. UMP,
P6. UMP *invariant*,
P7. UMP *unbiased*,
P8. *most stringent*,
P9. *locally minimax*,
P10. *asymptotically minimax*.

3. Basic theorems. In this section we prove the basic theorems that permit us to transform problems involving patterned covariance matrices into products of unpatterned problems. First we prove two linear algebra results. The first one tells how to transform the patterned covariance matrix to one that we can handle more easily. Before we state that theorem we introduce some notation. If A is a $k \times m$ matrix and $B = (b_{ij})$ is an $n \times p$ matrix, then the Kronecker product of A and B, $A * B$, is the $kn \times mp$ matrix

$$A * B = \begin{pmatrix} Ab_{11} & \cdots & Ab_{1p} \\ \vdots & & \vdots \\ Ab_{n1} & \cdots & Ab_{np} \end{pmatrix}.$$

The only two properties of Kronecker products we need are

$$(3.1) \qquad (A * B)(C * D) = AC * BD \quad \text{and hence} \quad (A * B)^{-1} = A^{-1} * B^{-1}.$$

$$(3.2) \qquad (A * B)' = A' * B'.$$

Both these properties follow easily from the definition.

Let I_j be the $j \times j$ identity, E be the $k \times k$ matrix of 1's and F be the $k \times k$ matrix with $F_{11} = k$, $F_{ij} = 0$ for $i + j > 2$. Let f be $k^{-\frac{1}{2}}$ times the $k \times 1$ vector of 1's. That is $f = (k^{-\frac{1}{2}}, \cdots, k^{-\frac{1}{2}})'$. Let C be an orthogonal matrix whose first column is f. Then $C'EC = F$.

LEMMA. *Let Σ have pattern A_k and $\Gamma = I_p * C$. Then Γ is an orthogonal matrix and*

$$(3.3) \qquad \Gamma'\Sigma\Gamma = \begin{pmatrix} \Sigma_1 + (k-1)\Sigma_2 & 0 \\ 0 & (\Sigma_1 - \Sigma_2) * I_{k-1} \end{pmatrix}.$$

PROOF. $\Gamma' = (I_p * C)' = (I_p * C') = (I_p * C^{-1}) = \Gamma^{-1}$, so Γ is orthogonal. If Σ has pattern A_k then

$$\Sigma = (\Sigma_1 - \Sigma_2) * I_k + \Sigma_2 * E .$$

Therefore

$$\begin{aligned}\Gamma'\Sigma\Gamma &= (I_p * C')[((\Sigma_1 - \Sigma_2) * I_k) + (\Sigma_2 * E)](I_p * C) \\ &= ((\Sigma_1 - \Sigma_2) * C'C) + (\Sigma_2 * C'EC) = [(\Sigma_1 - \Sigma_2) * I_k] + [\Sigma_2 * F] \\ &= \begin{pmatrix} \Sigma_1 + (k-1)\Sigma_2 & 0 \\ 0 & \Sigma_1 - \Sigma_2 * I_{k-1} \end{pmatrix} . \qquad \square\end{aligned}$$

Let

$$\Xi_1 = \Sigma_1 + (k-1)\Sigma_2 , \qquad \Xi_2 = \Sigma_1 - \Sigma_2 .$$

The transformation from Σ_1 and Σ_2 to Ξ_1 and Ξ_2 is invertible so we can consider Ξ_1 and Ξ_2 a reparametrization of Σ_1 and Σ_2.

The second lemma tells us what happens to the means when we use Γ to transform the covariances. If X is a $kp \times r$ matrix $X' = (X_1', \cdots, X_k')$, where X_i is $p \times r$, then we define $t(X)$ to be the $p \times rk$ matrix

$$t(X) = (X_1 \cdots X_k) .$$

$t(X)$ just rearranges the variables and serves as a bookkeeping device. Together with the following lemma, it prevents repetition of the same argument.

LEMMA. *If A is a $k \times k$ matrix, then*

$$t((I_p * A)X) = t(X)(I_r * A') .$$

PROOF. A direct computation yields

$$t[(I_p * A)X] = (\textstyle\sum_{j=1}^k a_{1j}X_j, \cdots, \sum_{j=1}^k a_{kj}X_j) = t(X)(I_r * A') . \qquad \square$$

Now we are ready for the basic theorems. Suppose

$$X_i(pk \times r_i) \sim N(\mu_i, a_i\Sigma) , \qquad i = 1, \cdots, k, \quad S \sim W(n, \Sigma) \tag{3.4}$$

where the X_i's and S are all mutually independent (this framework is just general enough to include canonical forms of the classification and multivariate analysis of variance problems). Let

$$\Gamma'S\Gamma = T = (T_{ij}) , \qquad \Gamma'X_i = U_i , \tag{3.5}$$

where T_{ij} is a $p \times p$ matrix for $i, j = 1, \cdots, k$, where $\Gamma = I_p * C$ and C is an orthogonal matrix whose first column is f. Let

$$W = T_{11} , \qquad V = \textstyle\sum_{j=2}^k T_{jj} , \qquad (Z_i, Y_i) = t(U_i) = t(X_i)(I_r * C) \tag{3.6}$$

where Z_i is $p \times r_i$, Y_i is $p \times r_i(k-1)$.

THEOREM 1. *Let $X_1, \cdots, X_m$, and S be mutually independent and have the distributions given in* (3.4). *If V, W, $Y_1, \cdots, Y_m$, and $Z_1, \cdots, Z_m$ are defined by* (3.5) *and* (3.6), *then they are sufficient, mutually independent and have the following* *distributions*:

$$(3.7) \qquad Z_i(p \times r_i) \sim N(\delta_i, a_i \Xi_1)\,, \qquad W \sim W(n, \Xi_1)\,,$$
$$Y_i(p \times r_i(k-1)) \sim N(\nu_i, a_i \Xi_2)\,, \qquad V \sim W(n(k-1), \Xi_2)\,,$$

where $\Xi_1 = \Sigma_1 + (k-1))\Sigma_2$, $\Xi_2 = \Sigma_1 - \Sigma_2$, $(\delta_i, \nu_i) = t(\mu_i)(I_{r_i} * C)$.

PROOF. Since X and S are sufficient, so are T (defined in (3.5)), Y_i and Z_i, $i = 1, \cdots, k$.

$$T \sim W(n, \Xi)\,, \qquad \Xi = \begin{pmatrix} \Xi_1 & 0 \\ 0 & \Xi_2 * I_{k-1} \end{pmatrix}.$$

$$(3.8) \qquad \begin{aligned} \mathscr{L}(T) &= \Xi^{-n/2} f(t)\,\mathrm{etr}\,(-\tfrac{1}{2}\Xi^{-1}T) \\ &= \Xi^{-n/2} f(t)\,\mathrm{etr}\,(-\tfrac{1}{2}\Xi_1^{-1}T_{11} - \tfrac{1}{2}\Xi_2^{-1}\textstyle\sum_{j=2}^{k} T_{jj}) \\ &= f(T)\,\mathrm{etr}\,(-\tfrac{1}{2}\Xi_1^{-1}W - \tfrac{1}{2}\Xi_2^{-1}V)\,. \end{aligned}$$

So by the factorization theorem, W, V, Y_i and Z_i are sufficient. By (3.8) the T_{ii} are all independent, so that $W \sim W(n, \Xi_1)$, $V \sim W(n(k-1), \Xi_2)$ and W and V are independent. Let $Y_i = (Y_{i,1} \cdots Y_{i,k-1})$, then

$$\begin{pmatrix} Z_i \\ Y_{i1} \\ \vdots \\ Y_{i,k-1} \end{pmatrix} = U_i \sim N(\Gamma'\mu_i, \Xi)\,.$$

Therefore Z_i and Y_i are independent and $Z_i \sim N(\delta_i, a_i\Xi_1)$, $Y_i \sim N(\nu_i, a_i\Xi_2)$ where δ_i and ν_i are given by

$$(\delta_i, \nu_i) = E(t(X_i))(I_{r_i} * C) = t(\mu_i)I_{r_i} * C)\,. \qquad \square$$

Theorem 1 shows that we can transform random variables whose distributions have patterned covariance matrices into two independent random variables $(Z_1, \cdots, Z_m, W)$ and $(Y_1, \cdots, Y_m, V)$. The following theorem shows that for all the problems we consider there is no relation between $(\delta_1, \cdots, \delta_m, \Xi_1)$ and $(\nu_1, \cdots, \nu_m, \Xi_2)$. Therefore Theorems 1 and 2 show that we can transform problems involving patterned covariance matrices to products of unpatterned problems.

THEOREM 2. (a) $\Sigma > 0$ *if and only if* $\Xi_1 > 0$, $\Xi_2 > 0$.
(b) $\mu_i = 0$ *if and only if* $\delta_i = 0$, $\nu_i = 0$.
(c) $-\infty < \mu_i < \infty$ *if and only if* $-\infty < \delta_i < \infty$, $-\infty < \nu_i < \infty$.
(d) *If* $r_i = r_j$ *then* $\mu_i = \mu_j$ *if and only if* $\delta_i = \delta_j$, $\nu_i = \nu_j$.
(e) *The columns of* μ_i *have pattern* B_k *if and only if* $-\infty < \delta_i < \infty$, $\nu_i = 0$.

PROOF. (a) $\Sigma > 0$ if and only if

$$\Gamma'\Sigma\Gamma = \begin{pmatrix} \Xi_1 & 0 \\ 0 & \Xi_2 * I_{k-1} \end{pmatrix} > 0$$

if and only if $\Xi_1 > 0$, $\Xi_2 > 0$.

(b), (c), (d) follow from the non-singularity of C and Γ.

(e) Suppose the columns of μ_i have pattern B_k. Then $t(\mu_i) = \theta_i * f'$ for some θ_i, $p \times r_i$. Therefore

$$(\delta_i, \nu_i) = (\theta_i * f')(I_{r_i} * C) = \theta_i * (f'C) = (\theta_i, 0, \cdots, 0) .$$

That finishes the only if part.

Now suppose $\nu_i = 0$. Then

$$t(\mu_i) = (\delta_i, 0)(I_{r_i} * C') = (\delta_i, 0)\begin{pmatrix} I_{r_i} * f' \\ I_{r_i} * \beta' \end{pmatrix} . \qquad \square$$

4. Testing that μ is patterned when Σ is. In this section we consider the problem of testing that the mean of a multivariate normal distribution is patterned when we know that the covariance is. We have $X(pk \times 1) \sim N(\mu, \Sigma)$, $S \sim W(n, \Sigma)$ and we are testing that μ has pattern B_k when we know Σ has pattern A_k. That is, we are testing the problem P_1.

$$\begin{aligned} P_1\colon\ & X(pk \times 1) \sim N(\mu, \Sigma) , \qquad S \sim W(n, \Sigma) , \\ & H\colon\ \mu \text{ has pattern } B_k, \quad \Sigma \text{ has pattern } A_k, \quad \Sigma > 0 , \\ & A\colon\ -\infty < \mu < \infty , \quad \Sigma \text{ has pattern } A_k, \quad \Sigma > 0 . \end{aligned} \tag{4.1}$$

When $p = 1$, this problem is the intraclass correlation model of Wilks (1946) and for $p > 1$ is considered in Olkin (1970).

THEOREM 3. *The problem P_1 defined in* (4.1) *can be transformed to the product of the trivial problem P_1' and the multivariate analysis of variance* (MANOVA) *problem P_1''*

$$\begin{aligned} P_1'\colon\ & Z(p \times 1) \sim N(\delta, \Xi_1) , & P_1''\colon\ & Y(p \times (k-1)) \sim N(\nu, \Xi_2) , \\ & W \sim W(n, \Xi_1) , & & V \sim W(n(k-1), \Xi_2) \\ & H\colon\ -\infty < \delta < \infty, \Xi_1 > 0 , & & H\colon\ \nu = 0, \Xi_2 > 0 , \\ & A\colon\ -\infty < \delta < \infty, \Xi_1 > 0 , & & A\colon\ -\infty < \nu < \infty, \Xi_2 > 0 . \end{aligned}$$

PROOF. By Theorem 1, Z, Y, W and V defined in (3.5) and (3.6) are sufficient and have the given distributions. By Theorem 2a and 2e, if Σ has pattern A_k, then μ has pattern B_k, $\Sigma > 0$ if and only if $-\infty < \delta < \infty$, $\nu = 0$, $\Xi_1 > 0$, and $\Xi_2 > 0$, and by Theorem 2(a) and 2(c), $-\infty < \mu < \infty$, $\Sigma > 0$ if and only if $-\infty < \delta < \infty$, $-\infty < \nu < \infty$, $\Xi_1 > 0$ and $\Xi_2 > 0$. $\square$

So P_1 is the product of a trivial problem and a MANOVA problem. By Theorem B any optimal procedure for the MANOVA problem is an optimal procedure for P_1. Invariance presents the only possible difficulty. It is possible that the maximal invariant for P_1 reduced by the largest group leaving P_1 invariant may include elements from P_1'. The following lemma shows that does not happen. P_1' is invariant under G_1': $Z \to AZ + C$, $W \to AWA'$ where A is non-singular. P_1'' is invariant under $Y \to BY\Gamma$, $V \to BVB'$ where B is nonsingular and Γ is orthogonal.

LEMMA. *$P_1 = P_1' \times P_1''$ is invariant under $G_1 = G_1' \times G_1''$. A maximal invariant for this group is the set of roots of $Y'V^{-1}Y$.*

PROOF. By Theorem A(vi), P_1 is invariant under G_1. By Theorem A(vii), P_1/G_1 is the product of P_1'/G_1' and P_1''/G_1''. The roots of $Y'V^{-1}Y$ are a maximal invariant for P_1'' (Lehmann (1959) pages 296–298). The whole (Z, W) space is on one orbit for P_1' so the roots of $Y'V^{-1}Y$ are a maximal invariant for P_1. □

We have now reduced P_1 to a MANOVA problem and we can use theorems about that problem to get theorems about P_1. To give an idea of the power of Theorem 3 and Theorem B, we give a summary of some known results about the MANOVA problem that transfer to P_1. If $\lambda_1 > \lambda_2 \cdots \lambda_a$ are the nonzero roots of $Y'V^{-1}Y$, let

$$\begin{aligned} f_1 &= \textstyle\prod_{i=1}^a (1+\lambda_i)^{-1} = |I + Y'V^{-1}Y|^{-1}, \\ f_2 &= -\textstyle\sum_{i=1}^a \lambda_i = -\operatorname{tr} YY'V^{-1}, \\ f_3 &= -\textstyle\sum_{i=1}^a \frac{\lambda_i}{1+\lambda_i} = -\operatorname{tr} YY'(YY'+V)^{-1}, \\ f_4 &= -\lambda_1 . \end{aligned} \tag{4.2}$$

In the following summary, references for the results of the MANOVA problem are given in parenthesis. They all carry over to P_1 by Theorem B.

(i) The LRT function is $cf_1^{(n+1)(k-1)/2}$ for some $c > 0$ (Anderson (1958) pages 187–190).

(ii) $Cf_1^{(k-1)/2}$ and $\exp(f_3/2)$ are Bayes for some $C > 0$ (Kiefer and Schwartz (1965)).

(iii) f_1, f_2, f_3 and f_4 are admissible (Schwartz (1967a)).

(iv) f_1, f_2 and f_4 are unbiased and have monotone power in the roots of $\nu'\Xi_2^{-1}\nu$ (Das Gupta, Anderson and Mudholkar (1964)).

(v) f_3 is locally minimax as $r = \operatorname{tr} \nu'\Xi_2^{-1}\nu \to 0$ (Schwartz (1967b)).

(vi) If $p = 1$ (the intraclass correlation model of Wilks), then P_1'' is a univariate analysis of variance problem and we can conclude f_1 is UMP invariant, UMP unbiased and most stringent (Lehmann (1959) pages 266–269).

(vii) If $k = 2$ then P_1'' is the Hotelling's T^2 problem and f_1 is UMP invariant, locally and asymptotically minimax as $\nu'\Xi_2^{-1}\nu \to 0$ and ∞ (Giri and Kiefer (1964)), and in some simple cases most stringent (Giri, Kiefer and Stein (1963)).

5. Multivariate analysis of variance. In this section we consider the multivariate analysis of variance (MANOVA) problem when, in addition to the usual assumptions, we assume that the covariance matrix has pattern A_k. A canonical form for the MANOVA problem is the following (see Lehmann (1959) pages 293–296.). We have a $p \times r$ random matrix X_1 and a $p \times s$ random matrix X_2 such that the columns of X_1 and X_2 are all independent, normally distributed, have common covariance matrix Σ and $EX = \mu_1$, $EY = \mu_2$, and a $p \times p$ matrix S that has a Wishart distribution with n degrees of freedom and $ES = n\Sigma$. We are testing $\mu_1 = 0$ versus $-\infty < \mu_1 < \infty$. That is, the MANOVA problem is the problem

$$\begin{aligned} Q: & X_1(p \times r) \sim N(\mu_1, \Sigma), \quad X_2(p \times s) \sim N(\mu_2, \Sigma), \quad S \sim W(n, \Sigma), \\ H: & \mu_1 = 0, \quad -\infty < \mu_2 < \infty, \quad \Sigma > 0, \\ A: & -\infty < \mu_1 < \infty, \quad -\infty < \mu_2 < \infty, \quad \Sigma > 0. \end{aligned}$$

So in this section we consider the problem P_2

$$\begin{aligned} P_2: & X_1(pk \times r) \sim N(\mu_1, \Sigma), \quad X_2(pk \times s) \sim N(\mu_2, \Sigma), \quad S \sim W(n, \Sigma), \\ & H: \mu_1 = 0, \quad -\infty < \mu_2 < \infty, \quad \Sigma \text{ has pattern } A_k, \\ & \quad \Sigma > 0, \\ & A: -\infty < \mu_1 < \infty, \quad -\infty < \mu_2 < \infty, \\ & \quad \Sigma \text{ has pattern } A_k, \quad \Sigma > 0, \end{aligned} \tag{5.1}$$

THEOREM 4. *P_2 defined in* (5.1) *can be transformed into the product of the two* MANOVA *problems P_2' and P_2'',*

$$\begin{aligned} P_2': & Z_1(p \times r) \sim N(\delta_1, \Xi_1), \quad Z_2(p \times s) \sim N(\delta_2, \Xi_1), \quad W \sim W(n, \Xi_1), \\ H: & \delta_1 = 0, \quad -\infty < \delta_2 < \infty, \quad \Xi_1 > 0, \\ A: & -\infty < \delta_1 < \infty, \quad -\infty < \delta_2 < \infty, \quad \Xi_1 > 0, \\ P_2'': & Y_1(p \times r(k-1)) \sim N(\nu_1, \Xi_2), \quad Y_2(p \times s(k-1)) \sim N(\nu_2, \Xi_2), \\ & V \sim W(n(k-1), \Xi_2), \\ & H: \nu_1 = 0, \quad -\infty < \nu_2 < \infty, \quad \Xi_2 > 0, \\ & A: -\infty < \nu_1 < \infty, \quad -\infty < \nu_2 < \infty, \quad \Xi_2 > 0. \end{aligned}$$

PROOF. By Theorem 1, Z_1, Z_2, Y_1, Y_2, W and V (defined in (3.5) and (3.6)) are independent, sufficient and have the distributions shown. By Theorem 2 (a), (b) and (c) $\mu_1 = 0$, $-\infty < \mu_2 < \infty$, $\Sigma > 0$ if and only if $\delta_1 = 0$, $\nu_1 = 0$, $-\infty < \delta_2 < \infty$, $-\infty < \nu_2 < \infty$, $\Xi_1 > 0$, $\Xi_2 > 0$. By Theorem 2 (a) and (c) $-\infty < \mu_1 < \infty$, $-\infty < \mu_2 < \infty$, $\Sigma > 0$ if and only if $-\infty < \delta_1 < \infty$, $-\infty < \nu_1 < \infty$, $-\infty < \delta_2 < \infty$, $-\infty < \nu_2 < \infty$, $\Xi_1 > 0$, $\Xi_2 > 0$. □

P_2 is a product of problems where neither problem is trivial, so we cannot use Theorem B. Theorem A is not quite so easy to use and it does not give quite such powerful results. However, we can derive some interesting results.

First, we reduce P_2 by invariance. P_2' is invariant under G_2': $Z_1 \to AZ\Gamma$, $Z_2 \to AZ_2 + B$, $W \to AWA'$ where A is non-singular and Γ is orthogonal. P_2'' is invariant under G_2'': $Y_1 \to CY_1\beta$, $Y_2 \to CY_2 + D$, $V \to CVC'$ where C is non-singular and β is orthogonal.

LEMMA. *$P_2 = P_2' \times P_2''$ is invariant under $G_2 = G_2' \times G_2''$. A maximal invariant for P_2 is the set of roots of $Z_1'W^{-1}Z_1$ and $Y_1V^{-1}Y_1$.*

PROOF. P_2 is invariant G_2 by Theorem A(vi). A maximal invariant for P_2' is the set of roots of $Z_1'W^{-1}Z_1$ and a maximal invariant for P_2'' is the set of roots of $Y_1'V^{-1}Y_1$ (Lehmann (1959) pages 296–298). By Theorem A(vii), the two sets together are a maximal invariant for P_2. □

Therefore, let $a_1 = \min(p, r)$, $b_1 = \max(p, r)$, $a_2 = \min(p, r(k-1))$, $b_2 = \max(p, r(k-1))$. Let $(\lambda_{11}, \cdots, \lambda_{1a_1})$ be the roots of $Z_1' W_1^{-1} Z_1$ and $(\lambda_{21}, \cdots, \lambda_{2a_2})$ be the roots of $Y_1' W_2^{-1} Y_1$. We consider only two invariant test functions for P_2.

$$
\begin{aligned}
&f_1 = \textstyle\prod_{i=1}^{a_1} (1 + \lambda_{1i})^{-1}, \qquad f_2 = \prod_{i=1}^{a_2} (1 + \lambda_{2i})^{-1}, \\
&g_1 = -\textstyle\sum_{i=1}^{a_1} \lambda_{1i}/(1 + \lambda_{1i}), \qquad g_2 = -\sum_{i=1}^{a_2} \lambda_{2i}/(1 + \lambda_{2i}), \\
&f = f_1 f_2^{k-1}, \qquad g = g_1 + g_2.
\end{aligned} \tag{5.2}
$$

THEOREM 5. (a) $(\lambda_{11}, \cdots, \lambda_{1a_1}, \lambda_{21}, \cdots, \lambda_{2a_2})$ *is a maximal invariant.*

(b) $\Lambda = cf^{(n+1)r/2}$ *is the* LRT *function for some* $c > 0$.

(c) *Under the null hypothesis*

$$f \sim \textstyle\prod_{i=1}^{a_1} \beta_{1i} \prod_{j=1}^{a_2} \beta_{2j}^{k-1},$$

where

$$\beta_{1i} \sim \mathrm{Be}\,(0; (n - p + i)/2, b_1/2).$$
$$\beta_{2i} \sim \mathrm{Be}\,(0; (n(k-1) - p + i)/2, b_2/2),$$

(d) $Cf^{r/2}$ *and* $\exp(\frac{1}{2} g)$ *are Bayes for some* $C > 0$.

(e) f *and* g *are admissible.*

(f) f *is unbiased and has monotone power in the roots of* $\delta_1' \Xi_1^{-1} \delta_1$ *and* $\nu_1' \Xi_2^{-1} \nu_1$.

PROOF. (a) This is the lemma above.

(b) Anderson (1958) pages 187–190 shows that $c_1 f_1^{(n+1)r/2}$ and $c_2 f_2^{(n+1)(k-1)r/2}$ are the LRT functions for P_2' and P_2'' respectively. Therefore, by Theorem A(ii), $cf^{(n+1)r/2} = c_1 c_2 f_1^{(n+1)r/2} f_2^{(n+1)(k-1)r/2}$ is the LRT for P_2.

(c) Anderson (1958) pages 193–195 shows that under the null hypothesis

$$f_1 \sim \textstyle\prod_{i=1}^{a_1} \beta_{1i}, \qquad f_2 \sim \prod_{i=1}^{a_1} \beta_{2i}.$$

Therefore $f = f_1 f_2^{k-1}$ has the distribution shown.

(d) Kiefer and Schwartz (1965) show that $C_1 f^{r/2}$ and $\exp(g_1/2)$ are Bayes for P_2' and that $C_2 f_2^{r(k-1)/2}$ and $\exp(g_2/2)$ are Bayes for P_2''. Therefore, by Theorem A-3, $Cf^{r/2} = C_1 C_2 f_1^{r/2} f_2^{r(k-1)/2}$ and $\exp(g/2) = \exp(g_1/2) \times \exp(g_2/2)$ are Bayes for P_2.

(e) This follows from (d).

(f) Das Gupta, Anderson and Mudholkar (1964) show that f_1 has monotone power in the roots of $\delta_1' \Xi_1^{-1} \delta_1$ and f_2 has monotone power in the roots of $\nu_1' \Xi_2^{-1} \nu_1$. By Theorem A(v), therefore, $f = f_1 f_2^{k-1}$ has monotone power in the roots of $\delta_1' \Xi_1^{-1} \delta_1$ and $\nu_1 \Xi_2^{-1} \nu_1$ and is therefore unbiased. □

For a Box–Anderson approximation to the null distribution of f see Arnold (1970).

Geisser (1963) considers the Hotelling's T^2 problem when the covariance matrix is patterned. Since the Hotelling's T^2 problem is just a special case of the MANOVA problem (when $r = 1$, $s = 0$), Theorems 4 and 5 apply to that problem also. It is interesting that unless $k = 2$ the Hotelling's T^2 problem does not transform into a product of Hotelling's T^2 problems as we might expect (since for this problem Y has dimentions $p \times k - 1$).

6. Multivariate classification. In this section we consider the multivariate classification problem when the covariance matrix is patterned. We only consider the problem of classifying the observation into one of two populations, but the extension to k populations is obvious. In the two population classification problem we have X_0 normally distributed with mean μ_0 and covariance matrix Σ, $X_{11}, \cdots, X_{1N_1}$, normally distributed with mean μ_1 and covariance matrix Σ, and $X_{21}, \cdots, X_{2N_2}$ normally distributed with mean μ_2 and covariance matrix Σ and we are testing $\mu_0 = \mu_1$ versus $\mu_0 = \mu_2$. We can put this into a canonical form as follows. Let X_1 and X_2 denote the sample means and S the pooled cross product matrix, i.e.,

$$\begin{aligned} &X_1 = \textstyle\sum_{i=1}^{N_1} X_{1i}/N_1 , \qquad X_2 = \sum_{j=1}^{N_2} X_{2i}/N_2 , \\ &S = \textstyle\sum_{i=1}^{N_1} (X_{1i} - X_1)(X_{1i} - X_1)' + \sum_{i=1}^{N_2} (X_{2i} - X_2)(X_{2i} - X_2)' , \\ &n_1 = N_1 - 1 , \qquad n_2 = N_2 - 1 . \end{aligned}$$

Then the classification problem becomes

$$\begin{aligned} Q: \ &X_0(p \times 1) \sim N(\mu_0, \Sigma) , \qquad S \sim W(n_1 + n_2, \Sigma) , \\ &X_1(p \times 1) \sim N(\mu_1, \Sigma/N_1) , \qquad X_2(p \times 1) \sim N(\mu_2, \Sigma/N_2) , \\ &H: \ \mu_0 = \mu_1 , \qquad -\infty < \mu_2 < \infty , \qquad \Sigma > 0 , \\ &A: \ \mu_0 = \mu_2 , \qquad -\infty < \mu_1 < \infty , \qquad \Sigma > 0 . \end{aligned}$$

In this paper we make the additional assumption that Σ is patterned. So, we consider the problem P_3

$$\begin{aligned} P_3: \ &X_0(pk \times 1) \sim N(\mu_0, \Sigma) , \qquad S \sim W(n_1 + n_2, \Sigma) , \\ &X_1(pk \times 1) \sim N(\mu_1, \Sigma/N_1) , \qquad X_2(pk \times 1) \sim N(\mu_2, \Sigma/N_2) , \\ &H: \ \mu_0 = \mu_2 , \qquad -\infty < \mu_1 < \infty , \qquad \Sigma \text{ has pattern } A_k , \qquad \Sigma > 0 , \\ &A: \ \mu_0 = \mu_1 , \qquad -\infty < \mu_2 < \infty , \qquad \Sigma \text{ has pattern } A_k , \qquad \Sigma > 0 . \end{aligned} \tag{6.1}$$

As in Sections 4 and 5 we can use Theorems 1 and 2 to transform this problem into a product of unpatterned problem.

THEOREM 6. *P_3 given in* (6.1) *is equivalent to the product of the classification problem P_3' and the generalized classification problem P_3''*:

$$\begin{aligned} P_3': \ &Z_0(p \times 1) \sim N(\delta_0, \Xi_1) , \qquad W \sim W((n_1 + n_2), \Xi_1) , \\ &Z_1(p \times 1) \sim N(\delta_1, \Xi_1/N_1) , \qquad Z_2(p \times 1) \sim N(\delta_2, \Xi_1/N_2) , \\ &H: \ \delta_0 = \delta_1 , \qquad -\infty < \delta_2 < \infty , \qquad \Xi_1 > 0 , \\ &A: \ \delta_0 = \delta_2 , \qquad -\infty < \delta_1 < \infty , \qquad \Xi_1 > 0 , \end{aligned}$$

$$\begin{aligned} P_3'': \ &Y_0(p \times (k-1)) \sim N(\nu_0, \Xi_2) , \qquad V \sim W((n_1 + n_2)(k-1), \Xi_2) , \\ &Y_1(p \times (k-1)) \sim N(\nu_1, \Xi_2/N_1) , \qquad Y_2(p \times (k-1)) \sim N(\nu_2, \Xi_2/N_2) , \\ &H: \ \nu_0 = \nu_1 , \qquad -\infty < \nu_2 < \nu , \qquad \Xi_2 > 0 , \\ &A: \ \nu_0 = \nu_2 , \qquad -\infty < \nu_1 < \infty , \qquad \Xi_2 > 0 . \end{aligned}$$

PROOF. Let Z_i, Y_i, W and V be defined in (3.5) and (3.6). Then by Theorem 1, Z_i, Y_i, W and V are sufficient and have the given distributions. By Theorem 2 the hypotheses transform as shown. □

Unfortunately, there has not been too much work done in the multivariate classification problem. We look at only one test function. Let

$$(6.2)\qquad \begin{aligned} f_1 &= \frac{1 + (N_2/(N_2+1))(Z_2 - Z_0)'W_1^{-1}(Z_2 - Z_0)}{1 + (N_1/(N_1+1))(Z_1 - Z_0)'W_1^{-1}(Z_1 - Z_0)}, \\ f_2 &= \frac{|I + (N_2/(N_2+1))(Y_2 - Y_0)'W_2^{-1}(Y_2 - Y_0)|}{|I + (N_1/(N_1+1))(Y_1 - Y_0)'W_2^{-1}(Y_1 - Y_0)|}, \\ f &= f_1 f_2^{k-1}. \end{aligned}$$

THEOREM 7. *Let P_3 be the problem defined in* (6.1), *f be defined by* (6.2). *Then $cf^{(N_1+N_2+1)/2}$ is the* LRT *function for some $c > 0$ and $f^{\frac{1}{2}}$ is Bayes (and hence f is admissible) for P_2.*

PROOF. Anderson (1958) pages 141–142 shows that $c_1 f_1^{(N_1+N_2+1)/2}$ is the LRT function for P_3'. By an easy generalization of his argument, $c_2 f_2^{(N_1+N_2+1)(k-1)/2}$ is the LRT function for P_3''. Therefore, by Theorem A(ii), $cf^{(N_1+N_2+1)/2} = c_1 f_1^{(N_1+N_2+1)/2} c_2 f_2^{(N_1+N_2+1)(k-1)/2}$ is the LRT function for P_3. Kiefer and Schwartz (1965) show that $f_1^{\frac{1}{2}}$ and $f_2^{(k-1)/2}$ are Bayes for P_3' and P_3'' respectively. Therefore, by Theorem A(iii), $f^{\frac{1}{2}} = f_1^{\frac{1}{2}} f_2^{(k-1)/2}$ is Bayes for P_3. □

When a problem P is the product of two non-trivial problems P_1 and P_2, it is often unclear how to use a "good" procedure for P_1 and a "good" procedure for P_2 to find a "good" procedure for P. Theorems A(ii) and A(iii), however, tell us how to use LRT and Bayes test functions for P_1 and P_2. It is interesting that for both the MANOVA and classification problem, these two approaches lead to functions that are equivalent.

7. Problems where the means and covariance matrices are both patterned. In this section, we shift our attention to problems where the means and covariance matrices are both patterned, i.e., problems involving interchangeable random variables (see Section 1). We prove a theorem showing that a general problem involving patterned means and patterned covariance matrices can be transformed to a product of a trivial problem and a problem identical to the original problem except nothing is patterned. Then we apply this theorem to the MANOVA and classification problems.

THEOREM 8. *The problem P*

$$\begin{aligned} P&: X(pk \times r) \sim N(\mu, \Sigma), \qquad S \sim W(n, \Sigma) \\ H&: \mu'\Sigma^{-1}\mu \in C, \quad \textit{the columns of } \mu \textit{ have pattern } B_k,\ \Sigma \textit{ has pattern } A_k \\ A&: \mu'\Sigma^{-1}\mu \in D, \quad \textit{the columns of } \mu \textit{ have pattern } B_k,\ \Sigma \textit{ has pattern } A_k \end{aligned}$$

can be transformed to a product of P'.

$$P': Z(p \times r) \sim N(\delta_1, \Xi_1), \qquad W \sim W(n, \Xi_1),$$
$$H': \delta'\Xi_1^{-1}\delta \in C,$$
$$A': \delta'\Xi_1^{-1}\delta \in D,$$

and the trivial problem P''

$$P'': Y(p \times (k-1) \sim N(\nu, \Xi_2), \qquad V \sim W(n(k-1), \Xi_2),$$
$$H'': \nu = 0, \qquad \Xi_2 > 0,$$
$$A'': \nu = 0, \qquad \Xi_2 > 0.$$

PROOF. Let W, V, Y and Z be given by (3.5) and (3.6). Then by Theorem Z, Y, W and V are sufficient and have the distributions shown. When μ is patterned, by Theorem 2(e),

$$\mu'\Sigma^{-1}\mu = (\Gamma'\mu)'(\Gamma'\Sigma\Gamma)^{-1}\Gamma'\mu$$
$$= (\delta', 0)\begin{pmatrix}\Xi_1^{-1} & 0 \\ 0 & \Xi_2^{-1} * I_{k-1}\end{pmatrix}\begin{pmatrix}\delta \\ 0\end{pmatrix} = \delta'\Xi_1^{-1}\delta. \qquad \square$$

P'' is invariant under the group G'': $Y \to AY\Gamma$, $W_2 \to AW_2A'$, where A is non-singular and Γ is orthogonal.

COROLLARY. *If P' is invariant under G', then P is invariant under $G = G' \times G''$ and a maximal invariant for P' under G' is a maximal invariant for P under G.*

PROOF. By Theorem A(vi), P is invariant under G. By Theorem A(vii), P/G is the product of P'/G' and P''/G''. A parameter maximal invariant for P'' is the set of roots of $\nu'\Xi_2^{-1}\nu$, which are 0 under both hypotheses. Therefore P''/G'' is a trivial simple problem, and by Theorem B, any sufficient statistic for P'/G' is a sufficient statistic for $P'/G' \times P''/G''$. $\square$

Theorem 8, its corollary and Theorem B imply that if we have an optimal procedure for a problem where nothing is patterned, we have an optimal procedure for the same problem when the means and covariances are patterned. As examples, we look at the MANOVA and multivariate classification problems.

A canonical form for the MANOVA problem (see Section 5) is the following.

$$Q: X_1(p \times r) \sim N(\mu_1, \Sigma), \qquad X_2(p \times s) \sim N(\mu_2, \Sigma), \qquad S \sim W(n, \Sigma)$$
$$H: \mu_1 = 0, \qquad -\infty < \mu_2 < \infty, \qquad \Sigma > 0,$$
$$A: -\infty < \mu_1 < \infty, \qquad -\infty < \nu_2 < \infty, \qquad \Sigma > 0.$$

Let $X = (X_1, X_2)$, $\mu = (\mu_1, \mu_2)$. Then $\mu_1 = 0$ if and only if

$$\mu'\Sigma^{-1}\mu = \begin{pmatrix}\mu_1'\Sigma^{-1}\mu_1, & \mu_1'\Sigma^{-1}\mu_2 \\ \mu_2'\Sigma^{-1}\mu_1, & \mu_2'\Sigma^{-1}\mu_2\end{pmatrix} = \begin{pmatrix}0 & 0 \\ 0 & \mu_2'\Sigma^{-1}\mu_2\end{pmatrix}. \tag{7.1}$$

So let A be the set of $(r + s) \times (r + s)$ positive semi-definite matrices having the pattern given in (7.1). Let B be the set of positive semi-definite matrices.

Then Q is problem

$$Q: X(p \times (r+s)) \sim N(\mu, \Sigma), \qquad S \sim W(n, \Sigma),$$
$$H: \ \mu'\Sigma^{-1}\mu \in A,$$
$$A: \ \mu'\Sigma^{-1}\mu \in B.$$

This is in the form of Theorem 8. Therefore, the MANOVA problem when the means and covariance matrices are patterned is the product of a trivial problem and an unpatterned MANOVA problem. When $r = 1$, $s = 0$, the MANOVA problem is the Hotelling's T^2 problem, and so the Hotelling's T^2 problem when the means and covariance matrix are both patterned is the product of a trivial problem and an unpatterned Hotelling's T^2 problem. In Section 4 we indicate that when we transform a problem to a product where one problem is trivial and the other is a problem that has been studied, then we have solved the product problem. So the Hotelling's T^2 problem and the MANOVA problem, when both means and covariance matrices are patterned, are both solved.

A canonical form for the classification problem (see Section 6) is the following.

$$R: X_0(p \times 1) \sim N(\mu_0, \Sigma), \qquad S \sim W(n_1 + n_2, \Sigma),$$
$$X_1(p \times 1) \sim N(\mu_1, \Sigma/N_1), \qquad X_2(p \times 1) \sim N(\mu_2, \Sigma/N_2),$$
$$H: \ \mu_0 = \mu_1, \qquad -\infty < \mu_2 < \infty, \qquad \Sigma > 0,$$
$$A: \ \mu_0 = \mu_2, \qquad -\infty < \mu_1 < \infty, \qquad \Sigma > 0.$$

To put this in the form of Theorem 8, let $X = (X_0, N_1^{\frac{1}{2}}X_1, N_2^{\frac{1}{2}}X_2)$ $\mu = (\mu_0, N_1^{\frac{1}{2}}\mu_1, N_2^{\frac{1}{2}}\mu_2)$. Then $\mu_0 = \mu_1$ if and only if

$$\mu'\Sigma^{-1}\mu = \begin{pmatrix} a & N_1^{\frac{1}{2}}a & b \\ N_1^{\frac{1}{2}}a & N_1 a & c \\ b & c & d \end{pmatrix} \tag{7.2}$$

for some a, b, c, d. So let A be the set of 3×3 positive semi-definite matrices having the form of (7.2).

Similarly $\mu_0 = \mu_1$ if and only if

$$\mu'\Sigma^{-1}\mu = \begin{pmatrix} a & b & N_2^{\frac{1}{2}}a \\ b & c & d \\ N_2^{\frac{1}{2}}a & d & N_2 a \end{pmatrix} \tag{7.3}$$

for some a, b, c, d. Let B be the set of 3×3 positive semi-definite matrices having the form of (7.3). Then R becomes

$$R: X(p \times 3) \sim N(\mu, \Sigma), \qquad S \sim W(n_1 + n_2, \Sigma)$$
$$H: \ \mu'\Sigma^{-1}\mu \in A$$
$$A: \ \mu'\Sigma^{-1}\mu \in B.$$

This is in the form for Theorem 8 and therefore the patterned classification problem is "solved" in the sense that every result about the unpatterned problem carries over to the patterned problem.

8. Other problems. In this section we give a short discussion of some other testing problems where we assume Σ has pattern A_k. Sections 4–7 give some idea of the results that are possible using Theorems A and B of Section 2 on products of problems. Most of the results of these sections follow easily from these results and Theorems 1 and 2.

Most other problems involving covariance matrices having pattern A_k can also be factored into products of standard problems. An example is the problem of testing that the means and covariance matrices of two normal distributions are the same. This is the problem P_1,

$$
\begin{aligned}
&P_1: X_1(pk \times 1) \sim N(\mu_1, \Sigma_1)\,, \qquad X_2(pk \times 1) \sim N(\mu_2, \Sigma_2)\,,\\
&S_1 \sim W(n_1, \Sigma_1)\,, \qquad S_2 \sim W(n_2, \Sigma_2)\,,\\
&H\colon\ \mu_1 = \mu_2\,, \qquad \Sigma_1 = \Sigma_2 > 0\,, \qquad \Sigma_1 \text{ and } \Sigma_2 \text{ have pattern } A_k\,,\\
&A\colon\ -\infty < \mu_1 < \infty\,, \qquad -\infty < \mu_2 < \infty\,, \qquad \Sigma_1 > 0\,, \qquad \Sigma_2 > 0\,,\\
&\qquad \Sigma_1 \text{ and } \Sigma_2 \text{ have pattern } A_k\,.
\end{aligned}
\tag{8.1}
$$

We can generalize Theorem 1 to show that P_1 is the product of P_1' and P_1''.

$$
\begin{aligned}
P_1'\colon\ & Z_1(p \times 1) \sim N(\delta_1, \Xi_{11})\,, \qquad Z_2(p \times 1) \sim N(\delta_2, \Xi_{12})\,,\\
& W_1 \sim W(n_1, \Sigma_{11})\,, \qquad W_2 \sim W(n_2, \Xi_{12})\,,\\
& H\colon\ \delta_1 = \delta_2\,, \qquad \Xi_{11} = \Xi_{12} > 0\,,\\
& A\colon\ -\infty < \delta_1 < \infty\,, \qquad -\infty < \delta_2 < \infty\,, \qquad \Xi_{11} > 0\,, \qquad \Sigma_{12} > 0\,,
\end{aligned}
$$

$$
\begin{aligned}
P_1''\colon\ & Y_1(p \times (k-1)) \sim N(\nu_1, \Xi_{21})\,, \qquad Y_2(p \times (k-1)) \sim N(\nu_2, \Xi_{22})\,,\\
& V_1 \sim W(n_1(k-1), \Sigma_{21})\,, \qquad V_2 \sim W(n_2(k-1), \Xi_{22})\,,\\
& H\colon\ \nu_1 = \nu_2\,, \qquad \Xi_{21} = \Xi_{22} > 0\,,\\
& A\colon\ -\infty < \nu_1 < \infty\,, \qquad -\infty < \nu_2 < \infty\,, \qquad \Xi_{21} > 0\,, \qquad \Xi_{22} > 0\,.
\end{aligned}
$$

The problem of testing the equality of two covariance matrices and the multivariate Behrens–Fisher problem, when Σ has pattern A_k, factor in the same way. In P_1, if we also assume that μ_1 and μ_2 have pattern B_k, then P_1 is the product of P_1' (a problem identical to P except nothing is patterned) and the non-trivial problem of testing the equality of two covariance matrices (since when the means are patterned, $\nu_1 = \nu_2 = 0$). So, not all problems where the means and covariances are both patterned transform to a product where one problem is trivial.

An example of a problem which does not factor is the following. We have $X' = (X_1', \cdots, X_k')$, where X_i is $p \times 1$. We want to test whether the X_i are independent when we know that Σ has pattern A_k (or perhaps when we know that the X_i are interchangeable, i.e., μ has pattern B_k aud Σ has pattern A_k). If X is a kp-variate normal distribution, then the X_i are independent if and only if Σ has pattern A_k (with submatrices Σ_1 and Σ_2) and $\Sigma_2 = 0$. This leads to the problem

$$
\begin{aligned}
Q: X(pk \times 1) \sim N(\mu, \Sigma)\,, \quad & S \sim W(n, \Sigma)\,, \\
H: \ -\infty < \mu < \infty\,, \quad & \Sigma \text{ has pattern } A_k\,, \quad \Sigma_2 = 0\,, \quad \Sigma_1 > 0\,, \\
A: \ -\infty < \mu < \infty\,, \quad & \Sigma \text{ has pattern } A_k\,, \quad \Sigma > 0\,.
\end{aligned}
$$

Let $\Xi_1 = \Sigma_1 + (k-1)\Sigma_2$, $\Xi_2 = \Sigma_1 - \Sigma_2$. Then $\Sigma_2 = 0$ if and only if $\Xi_1 = \Xi_2$. So Q transforms to

$$
\begin{aligned}
Q': \ & Z(p \times 1) \sim N(\delta, \Xi_1)\,, \quad X(p \times (k-1) \sim N(\nu, \Xi_2)\,, \\
& W \sim W(n, \Xi_1)\,, \quad V \sim W(n(k-1)\,, \Xi_2)\,, \\
& H: \ -\infty < \delta < \infty\,, \quad -\infty < \nu < \infty\,, \quad \Xi_1 = \Xi_2 > 0\,, \\
& A: \ -\infty < \delta < \infty\,, \quad -\infty < \nu < \infty\,, \quad \Xi_1 > 0\,, \quad \Xi_2 > 0\,.
\end{aligned}
$$

Q' is the problem of testing the equality of two covariance matrices, a problem that has been studied. Any results for Q' then yield results for Q.

Problems like Q that do not factor when we apply Theorems 1 and 2 are rare. Usually a problem P in which we assume Σ has pattern A_k factors into the product of a problem P_1 similar to P (except the covariance is no longer assumed patterned) and a generalized form of P_1 (where the p-variate normal random variable is replaced by a $(p \times (k-1))$-variate normal random variable).

REFERENCES

ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

ARNOLD, S. F. (1970). Products of problems and patterned covariance matrices that arise from interchangeable random variables. Stanford Univ. Technical Report No. 46.

DAS GUPTA, S., ANDERSON, T. W. and MUDHOLKAR, G. S. (1964). Monotonicity of the power function of some tests of the multivariate linear hypothesis. *Ann. Math. Statist.* **35** 200–205.

GEISSER, SEYMOUR (1963). Multivariate analysis of variance for a special covariance case. *J. Amer. Statist. Assoc.* **58** 660–669.

GIRI, N. and KIEFER, J. (1964). Local and asymptotic minimax properties of multivariate tests. *Ann. Math. Statist.* **35** 21–35.

GIRI, N., KIEFER, J. and STEIN, C. (1963). Minimax characteristic of Hotelling's T^2 in the simplest case. *Ann. Math. Statist.* **34** 1524–1535.

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. Statist.* **36** 747–770.

KRISHNAIAH, P. R. and PATHAK, P. K. (1967). Test for the equality of covariance matrices under the intraclass correlation model. *Ann. Math. Statist.* **38** 1286–1288.

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

OLKIN, INGRAM (1970). Statistical inference for patterns of symmetry. To appear as a technical report, Stanford Univ.

SCHWARTZ, RICHARD (1967a). Admissible tests in multivariate analysis. *Ann. Math. Statist.* **38** 698–710.

SCHWARTZ, RICHARD (1967b). Locally minimax tests. *Ann. Math. Statist.* **38** 340–360.

SRIVASTAVA, M. S. (1965). Some tests for the intraclass correlation model. *Ann. Math. Statist.* **36** 1802–1806.

WILKS, S. S. (1946). Simple criteria for testing equality of means, equality of variances and equality of covariances in a normal multivariate sample. *Ann. Math. Statist.* **17** 257–281.

DEPARTMENT OF MATHEMATICS
LAWRENCE UNIVERSITY
P. O. BOX 1847
APPLETON, WISCONSIN 54911