AD-735454

SOME INFERENCE PROBLEMS ASSOCIATED WITH THE COMPLEX MULTIVARIATE NORMAL DISTRIBUTION

by

John C. Young

Technical Report No. 102
Department of Statistics ONR Contract

April 1, 1971

Research sponsored by the Office of Naval Research
Contract N00014-68-A-0515
Project NR 042-260

DEPARTMENT OF STATISTICS
Southern Methodist University

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) SOUTHERN METHODIST UNIVERSITY	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP UNCLASSIFIED

3. REPORT TITLE

"Some Inference Problems Associated with the Complex Multivariate Normal Distribution"

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

John C. Young

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
April 1, 1971	93	21

8a. CONTRACT OR GRANT NO. N00014-68-A-0515 b. PROJECT NO. NR 042-260 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 102 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Office of Naval Research

13. ABSTRACT

Consider k = q+1 complex multivariate normal populations with the same variance-covariance matrix but with different means. The problem of the number of discriminant functions needed to discriminate among the k-populations is considered and shown to be equivalent to the rank of the mean-space. A test for this dimensionality being r is presented. Also developed is the statistic for testing the goodness of fit of a single hypothetical discriminant function. As in the case of the real normal populations, the test statistic for this single hypothetical function is presented as the product of two independent factors, whose distributions are given; one measuring the direction aspect of the hypothetical function, and the other measures the collinearity aspect that is the necessity of only one function.

Also included is the Bartlett decomposition of a complex Wishart matrix and some results pertaining to the coherence of complex random variables.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED
Security Classification

TABLE OF CONTENTS

CHAPTER I

1. Introduction

R. A. Wooding (1956) derived the p-variate complex normal distribution. Let $\underline{\xi}$ denote a complex normal random vector with components ξ_j $(j = 1, 2, \cdots, p)$. Then each ξ_j $(j = 1, 2, \cdots, p)$ represents a complex random variable with real and imaginary components given by x_j and y_j. It was found that if the covariance relation between the components of $\underline{\xi}$ had the following form

$$E(x_m \; x_n) = E(y_m \; y_n) \; ; \; E(x_m \; y_n) = -E(x_n \; y_m) \qquad (1.1.0)$$

then the p-variate complex normal density function could be represented as

$$\Pi^{-p}|L|^{-1} \exp(-\,\underline{\xi}^{*\prime}L^{-1}\underline{\xi}) \qquad (1.1.1)$$

where L is the non-singular Hermitian variance-covariance matrix and $\underline{\xi}^{*\prime}$ represents the transpose of the conjugate of $\underline{\xi}$. The mean of the vector $\underline{\xi}$ is assumed to be $\underline{0}$. Wooding (1956) also derived the characteristic function for the above distribution and found it to be given by

$$\exp\left(-\,\frac{1}{4}\,\underline{T}^{*\prime}L\underline{T}\right)$$

where $\underline{T}$ is some arbitrary complex parameter.

It appears that N. R. Goodman (1963), working independently and using the same assumptions as given in (1.1.0), has also derived the complex multivariate normal distribution. He extended the development further by deriving the complex analogue for the real Wishart distribution, denoting it as the complex Wishart distribution. Also, Goodman derived the distributions of the complex analogues for the sample multiple and sample partial correlation coefficients. In the complex case, these are denoted as the sample multiple coherence and the sample partial coherence.

All developments that have followed in this area of complex multivariate analysis have been based on the assumption of the density function specified by Wooding (1956) or Goodman (1963).

Other authors who made early contributions in this area of complex multivariate analysis have been M. S. Srivastava (1965), N. Giri (1965), D. G. Kabe (1966[b]), and C. G. Khatri (1965). D. G. Kabe (1965[a]) gave a simplified method of deriving the distribution of the sample multiple coherence and the sample partial coherence. He represented the density functions in a finite series while Goodman (1963) had presented these as infinite series. Kabe (1966[b]) developed the complex analogue some results in classical multivariate normal regression theory. Srivastava (1965) gives a direct and simplified method of deriving the Wishart distribution. Khatri's (1965) contributions were in many areas. He obtained the maximum likelihood estimates of the parameters of the complex multivariate normal given by Wooding (1956) and Goodman (1963); he then derived the distributions of these estimates. Also included in his work are procedures for transforming from densities of complex random variables to other complex random variables. Using these techniques, Khatri (1965) derived the distributions of the sample multiple coherence.

Khatri (1965) also derives the distribution of Wilks' Λ in the complex case. He gives a representation of Wilks' Λ (in the complex case) as the product of real independent beta variables. It is this development that proves useful to the work given in this dissertation.

Through the work of the above authors the foundation for complex multivariate analysis, based on the assumption of a complex multivariate density as given in (1.1.1), has been established. It is the purpose of this dissertation to extend this development by examining Wilks' Λ which appears as a solution to some inference problems associated with this particular complex multivariate normal distribution.

This extension is covered in Chapter III by an examination of Wilks' Λ as presented in the multivariate analysis of variance of complex normal data, the regression analysis of complex normal variables upon real dummy variables, and discrimination among several complex normal populations with the same variance-covariance matrix but with different means. Also the goodness of fit of a hypothetical discriminant function is considered.

In Chapter II, some basic theorems dealing with the complex multivariate normal are established along with the Bartlett decomposition of a complex Wishart matrix. These results are used in establishing some different expressions for the sample coherence and sample multiple coherence. These results prove useful in working with Wilks' Λ in Chapter III. Chapter IV presents a summary of the work covered in this dissertation along with possible extensions of this research into other areas.

2. Definitions and Notation

Let $\underline{\xi}$ denote a $p \times 1$ complex random vector,

$$\underline{\xi}' = [\xi_1, \xi_2, \cdots, \xi_p]_{1\times p} \qquad (1.2.0)$$

where $\xi_j = x_j + iy_j$. Denote the $(2p \times 1)$ vector of real and imaginary components by

$$\underline{\eta}' = [x_1, y_1, x_2, y_2, \cdots, x_p, y_p]_{1\times 2p} \tag{1.2.1}$$

and let this random vector have a 2p-variate normal distribution with

$$\underline{\delta}' = E(\underline{\eta}') = [\mu_1, \nu_1, \mu_2, \nu_2, \cdots, \mu_p, \nu_p]$$

and the variance-covariance matrix given by

$$\Sigma_{\underline{\eta}} = E[(\underline{\eta} - \underline{\delta})(\underline{\eta} - \underline{\delta})']$$

where the 2×2 submatrices of $\Sigma_{\underline{\eta}}$ have the following special form

$$\begin{bmatrix} E[x_j - E(x_j)][x_k - E(x_k)] & E[x_j - E(x_j)][y_k - E(y_k)] \\ E[y_j - E(y_j)][x_k - E(x_k)] & E[y_j - E(y_j)][y_k - E(y_k)] \end{bmatrix}$$

$$= \begin{cases} \dfrac{\sigma_k^2}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \text{if } j = k \\ \dfrac{\sigma_j \sigma_k}{2} \begin{bmatrix} \alpha_{jk} & -\beta_{jk} \\ \beta_{jk} & \alpha_{jk} \end{bmatrix} & \text{if } j \neq k \end{cases} \tag{1.2.2}$$

Then the density function of $\underline{\eta}$ is given by

$$P(\underline{\eta}) = \frac{1}{(2\pi)^p |\Sigma_{\underline{\eta}}|^{1/2}} \exp\left(-\frac{1}{2}(\underline{\eta} - \underline{\delta})' \Sigma_{\underline{\eta}}^{-1} (\underline{\eta} - \underline{\delta}) \right) \tag{1.2.3}$$

provided that $|\Sigma_{\underline{\eta}}| \neq 0$.

It has been shown by Wooding (1956) and Goodman (1963) that if condition (1.2.2) is satisfied for $\underline{\eta}$, then $\underline{\xi}$ is distributed as a p-variate complex normal vector with probability density function given by

$$P(\underline{\xi}) = (\Pi^{p}|\Sigma_{\underline{\xi}}|)^{-1} \exp[-(\underline{\xi} - \underline{\mu})^{*\prime}\Sigma_{\underline{\xi}}^{-1}(\underline{\xi} - \underline{\mu})] \qquad (1.2.4)$$

where the mean $\underline{\mu}$ is given as

$$\underline{\mu}' = E(\underline{\xi}') = [\mu_1+i\nu_1, \mu_2+i\nu_2, \cdots, \mu_p+i\nu_p]$$

and the variance-covariance structure is given by the Hermitian positive definite matrix $\Sigma_{\underline{\xi}}$, where

$$\Sigma_{\underline{\xi}} = E[(\underline{\xi} - \underline{\mu})(\underline{\xi} - \underline{\mu})^{*\prime}]$$

and the elements of $\Sigma_{\underline{\xi}}$ are given as

$$\sigma_{jk} = \begin{cases} \sigma_k^2 & \text{if } j = k \\ (\alpha_{jk} + i\beta_{jk})_{j\,k} & \text{if } j \neq k \end{cases} .$$

The density function (1.2.4) will be denoted as $CN_p(\underline{\mu}, \Sigma)$.

Let $F(\underline{\xi})$ denote the distribution function of the complex random vector $\underline{\xi}$ with density given as $CN_p(\underline{\mu}, \Sigma)$. By the probability that $\underline{\xi}$ is less than $\underline{\xi}_0$, the following is meant.

$$F(\underline{\xi}_0) = P(\underline{\xi} < \underline{\xi}_0) = P(\underline{\xi}_R < \underline{\xi}_{0R} ; \underline{\xi}_I \leq \underline{\xi}_{I0}) \qquad (1.2.5)$$

where $\underline{\xi}_R$ denotes the real part of $\underline{\xi}$ and $\underline{\xi}_I$ denotes the real random variable attached to the imaginary quantity.

Let $\underline{\xi}_1, \underline{\xi}_2, \cdots, \underline{\xi}_n$ denote n independent and identically distributed p-variate complex normal vectors with density given by $CN_p(\underline{\mu}, \Sigma)$. Then the random matrix S, defined by

$$S = \sum_{i=1}^{n} (\underline{\xi}_i - \bar{\underline{\xi}})(\underline{\xi}_i - \bar{\underline{\xi}})^{*'}$$

where

$$\bar{\underline{\xi}} = \frac{1}{n} \sum_{i=1}^{n} \underline{\xi}_i$$

has been shown by Goodman (1963) and Khatri (1965) to be distributed as a complex Wishart with probability density function given by

$$P(S) = [\Gamma_p(n)]^{-1} |\Sigma|^{-n} |S|^{n-1-p} \exp[- \text{tr } \Sigma^{-1} S]$$

where

$$\Gamma_p(n) = \pi^{\frac{1}{2}p(p-1)} \left\{ \prod_{j=1}^{p} \Gamma(n-j+1) \right\} .$$

The complex Wishart density function with parameter Σ and degrees of freedom n-1 will be denoted as $CW_p(S|n-1|\Sigma)$.

Let $\underline{\xi}$, where $\underline{\xi}' = [\xi_1, \xi_2, \cdots, \xi_p]$, be distributed as a $CN_p(\underline{\mu}, \Sigma)$. As in real multivariate analysis, the relationship between the component ξ_p and a linear combination of $\xi_1, \xi_2, \cdots, \xi_{p-1}$ has a very important role in complex multivariate analysis. In the real case this relationship is measured by the multiple correlation coefficient which is defined as the maximum correlation between one component and a linear combination of the others. In complex multivariate analysis, Goodman (1963) has termed this quantity as the multiple coherence and shown it to be given by

$$P^2_{\xi_p}(\xi_1, \xi_2, \cdots, \xi_{p-1}) = \frac{\Sigma^{*\prime}_{12}\Sigma_{11}\Sigma_{12}}{\sigma_{pp}} = \frac{\Sigma_{21}\Sigma_{11}\Sigma_{12}}{\sigma_{pp}} \qquad (1.2.6)$$

where Σ_{11}, Σ_{12}, Σ_{21} are obtained from the variance-covariance matrix Σ which is partitioned as

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \sigma_{pp} \end{array}\right] .$$

Goodman (1963) and Khatri (1965) have given the distribution of the sample estimate, $\hat{P}^2$, of $P^2_{\xi_p}(\xi_1, \xi_2, \cdots, \xi_{p-1})$ as

$$f(\hat{R}^2) = \frac{\Gamma(n)}{\Gamma(p-1)\Gamma(n-p\_1)} (1-P^2)^n (\hat{R}^2)^{p-2} (1-R^2)^{n-p} F(n, n ; p-1, P^2\hat{R}^2)$$

where P denotes (1.2.6), the population multiple coherence of ξ_p with $\xi_1, \xi_2, \cdots, \xi_p$. $\hat{R}^2$ denotes the sample estimate and $F(, ; ;)$ denotes the hypergeometric function, i.e.,

$$F(n, n ; p-1, P^2\hat{R}^2) = \sum_{j=0}^{\infty} \frac{\Gamma(n+j)}{\Gamma(n)} \cdot \frac{\Gamma(n+j)}{\Gamma(b)} \cdot \frac{\Gamma(p-1)}{\Gamma(p-1+j)} \cdot \frac{(P^2\hat{R}^2)^j}{j\,!} .$$

CHAPTER II

1. Basic Results

There is a great deal of similarity between the multivariate analysis of complex random variables and real random variables. Despite this similarity, it seems worthwhile to include the following results with formal proofs in order to provide a firm foundation for the material presented in latter chapters.

In the case of real normally distributed random variables, it is of great importance to establish the distribution of the square of the random variable. The analogy of this in the complex case is the complex random variable times its conjugate. Let $Z \sim CN_1(0, 1)$. As pointed out in Chapter I, Z is composed of two real, independently distributed normal variables with $\mu = 0$, and $\sigma^2 = 1/2$. Consider the real random variable defined by $y = ZZ^* = x^2 + y^2$ which is the sum of squares of two independent $N(0, 1/2)$ random variables. From real univariate theory, it is clear that

$$y = ZZ^* \sim \frac{\chi^2_{(2)}}{2} \quad . \qquad (2.1.0)$$

where $\chi^2_{(f)}$ stands for a chi-square random variable with f degrees of freedom.

Let $\underline{\xi}_{p\times 1}$ denote a complex random vector with probability density function given by

$$P(\underline{\xi}') = (\Pi^p|\Sigma|)^{-1} \exp[-(\underline{\xi} - \underline{\mu})^{*\prime}\Sigma^{-1}(\underline{\xi} - \underline{\mu})] \quad .$$

Consider the transformation given by

$$\underset{\sim}{Y}_{r\times 1} = A_{r\times p}\, \underset{\sim}{\xi}_{p\times 1}$$

where A is a real or complex matrix of rank $r \le p$. If $r = p$, then the probability density function of the random vector $\underset{\sim}{Y}$ is given by

$$P(\underset{\sim}{Y}') = (\Pi^{p}|\Sigma|)^{-1} \exp[-\ (A^{-1}\underset{\sim}{Y} - \underset{\sim}{\mu})^{*'}\Sigma^{-1}(A^{-1}\underset{\sim}{Y} - \underset{\sim}{\mu})] \cdot |J|$$

where $J(x \to y)$ is given by Khatri (1965)

$$J = |A|^{-1}|A^{*'}|^{-1} \quad .$$

Thus the density function of $\underset{\sim}{Y}$ becomes

$$P(\underset{\sim}{Y}') = (\Pi^{p}|A\Sigma A^{*'}|)^{-1}\exp[-\ (\underset{\sim}{Y} - A\underset{\sim}{\mu})^{*'}(A\Sigma A^{*'})^{-1}(\underset{\sim}{Y} - A\underset{\sim}{\mu})] \quad .$$

This result is stated formally as the following theorem.

Theorem 2.1. If $\underset{\sim}{\xi}$ is distributed accordingly to $CN_p(\underset{\sim}{\mu}, \Sigma)$, then $\underset{\sim}{Y} = A\underset{\sim}{\xi}$ is distributed accordingly to $CN_p(A\underset{\sim}{\mu}, A\Sigma A^{*'})$, where A is a $p \times p$ real or complex matrix of rank p.

Now consider the transformation

$$\underset{\sim}{Y}_{r\times 1} = A_{r\times p}\, \underset{\sim}{\xi}_{p\times 1}$$

where the matrix A is of rank $r < p$. The mean of $\underset{\sim}{Y}$ is given by

$$E(\underset{\sim}{Y}) = E(A\underset{\sim}{\xi}) = AE(\underset{\sim}{\xi}) = A\underset{\sim}{\mu}$$

and the covariance is given by

$$\begin{aligned} E[\underset{\sim}{Y} - E(\underset{\sim}{Y})][\underset{\sim}{Y} - E(\underset{\sim}{Y})^{*'}] &= AE(\underset{\sim}{\xi} - \underset{\sim}{\mu})(\underset{\sim}{\xi} - \underset{\sim}{\mu})^{*'}A^{*'} \\ &= A\Sigma A^{*'} \quad . \end{aligned}$$

Since $A_{r \times p}$ is of rank r, it is known that there exists a matrix $C_{(p-r)\times p}$ such that the transformation

$$\begin{pmatrix} \underline{Y}_{r\times 1} \\ \underline{Z}_{(p-r)\times 1} \end{pmatrix} = \begin{pmatrix} A \\ C \end{pmatrix} \underline{\xi}$$

is a non-singular transformation. Then $\underline{Y}$ and $\underline{Z}$ will have a joint complex normal distribution and the marginal distribution of $\underline{Y}$ will be $CN_r(A\underline{\mu}, A\Sigma A^{*'})$. This result will be stated as the following theorem.

Theorem 2.2. If $\underline{\xi}$ is distributed accordingly to $CN_p(\underline{\mu}, \Sigma)$, then $\underline{Y} = A\,\underline{\xi}$ is distributed accordingly to $CN_r(A\underline{\mu}, A\Sigma A^{*'})$, where A is a $r \times p$ matrix of real or complex elements of rank $r \leq p$.

Now let $\underline{\xi}_{p\times 1} \sim CN_p(\underline{0}, I_{p\times p})$, and make the transformation

$$\underline{Y}_{p\times 1} = U_{p\times p}\,\underline{\xi}_{p\times 1}$$

where U is a unitary matrix, i.e., $U^{*'}U = UU^{*'} = I_p$. Then by Theorem 2.1, $\underline{Y}_{p\times 1} \sim CN_p(\underline{0}, I_p)$. It should be noted that the elements of U must be constants. If the elements are random variables, the transformation is a random unitary transformation and the distribution of $\underline{Y}$ is not necessarily that of a $CN_p(\underline{0}, I_p)$. In either case

$$\underline{Y}^{*'}\underline{Y} = \underline{\xi}^{*'}U^{*'}U\,\underline{\xi} = \underline{\xi}^{*'}\underline{\xi} \quad . \tag{2.1.1}$$

Thus, $\underline{\xi}^{*'}\underline{\xi}$ remains invariant under a unitary transformation. Now consider the transformation

$$\underline{Y}_{r\times 1} = A_{r\times p}\,\underline{\xi}_{p\times 1} \tag{2.1.2}$$

where the matrix A is of rank $r < p$ and has the additional property

$A_{r\times p}A^{*\prime}_{p\times r} = I_r$. This transformation will be referred to as a semi-unitary transformation. As in Theorem 2.2, it is known that the matrix A may be augmented with a matrix B such that the transformation

$$\begin{pmatrix} \underline{Y}_{r\times 1} \\ \underline{Z}_{(p-r)\times 1} \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} \underline{\xi}$$

is of full rank and furthermore it is unitary. Thus

$$\begin{pmatrix} \underline{Y}_{r\times 1} \\ \underline{Z}_{(p-r)\times 1} \end{pmatrix} \sim CN_p(\underline{0}, I_p)$$

and from the invariance property (2.1.1)

$$\xi_1^*\xi_1 + \xi_2^*\xi_2 + \cdots + \xi_p^*\xi_p = Y_1^*Y_1 + Y_2^*Y_2 + \cdots + Y_r^*Y_r + Z_{r+1}^*Z_{r+1} + \cdots + Z_p^*Z_p$$

and the variables $Y_1, Y_2, \cdots, Y_r, Z_{r+1}, Z_{r+2}, \cdots, Z_p$ are distributed as independent $CN_1(0, 1)$ as are the $\xi_1, \xi_2, \cdots, \xi_p$.

From (2.1.1) we have

$$\underline{Y}^{*\prime}\underline{Y} = Y_1^*Y_1 + Y_2^*Y_2 + \cdots + Y_r^*Y_r$$

and since $Y_1, Y_2, \cdots, Y_r$ for an independent set of $CN_1(0, 1)$, then from (2.1.0)

$$Y_1^*Y_1 + Y_2^*Y_2 + \cdots + Y_r^*Y_r \sim \frac{\chi^2_{2r}}{2} .$$

Also

$$\xi_1^*\xi_1 + \xi_2^*\xi_2 + \cdots + \xi_p^*\xi_p \sim \frac{\chi^2_{2p}}{2} .$$

Then

$$\xi_1^*\xi_1 + \xi_2^*\xi_2 + \cdots + \xi_p^*\xi_p - (Y_1^*Y_1 + \cdots + Y_r^*Y_r) \sim \frac{\chi^2_{2(p-r)}}{2} .$$

Since

$$Z_{r+1}^*Z_{r+1} + \cdots + Z_p^*Z_p = (\xi_1^*\xi_1 + \cdots + \xi_p^*\xi_p) - (Y_1^*Y_1 + \cdots + Y_r^*Y_r) .$$

This result is the complex analogue of Fisher's Lemma, and is formally stated as the following theorem.

Theorem 2.2. If $\underline{\xi}_{p\times 1} \sim CN_p(\underline{0}, I_p)$, then $\underline{\xi}^*\underline{\xi}' \sim \frac{\chi^2_{2p}}{2}$ and if $\underline{Y}_{r\times 1} = A\underline{\xi}$, where A is a semi-unitary transformation, then $\underline{Y}^{*\prime}\underline{Y} \sim \frac{\chi^2_{2r}}{2}$. Moreover, $\underline{Z}^{*\prime}\underline{Z} = \underline{\xi}^{*\prime}\underline{\xi} - \underline{Y}^{*\prime}\underline{Y} \sim \frac{\chi^2_{2(p-r)}}{2}$ and is independent of $\underline{Y}^{*\prime}\underline{Y}$.

Let $X = [\underline{\xi}_1, \underline{\xi}_2, \cdots , \underline{\xi}_n]$ where the $\underline{\xi}_i \sim$ Independent $CN_p(\underline{0}, \Sigma)$ and consider matrix of Hermitian forms, $XAX^{*\prime}$ where A is a real idempotent matrix of rank r. The distribution of $XAX^{*\prime}$ needs to be established. Since A is a symmetric matrix, the spectral decomposition of A may be used, A may be written as

$$A = \sum_{i=1}^{p} \lambda_i \underline{\beta}_i \underline{\beta}_i'$$

where the λ_i are the characteristic roots of A, and the $\underline{\beta}_i$ are the corresponding orthogonal characteristic vectors of A. Since A is idempotent of rank r, then $\lambda_i = 1$, for $i = 1, \cdots , r$ and $\lambda_j = 0$ for $j = r + 1, \cdots , p$, thus the spectral decomposition of A is given by

$$A = \sum_{i=1}^{r} \underline{\beta}_i \underline{\beta}_i'$$

and

$$XAX^{*\prime} = X \sum_{i=1}^{r} \underline{\beta}_i \underline{\beta}_i' X^{*\prime} = \sum_{i=1}^{r} X\underline{\beta}_i \underline{\beta}_i' X^{*\prime} \quad .$$

Since $\underline{\xi}_i \sim CN_p(\underline{0}, \Sigma)$, then the distribution of the matrix $X_{p\times n}$ is given by $CN_{p\times n}(\emptyset, I \otimes \Sigma)$ where $\otimes$ stands for the Kronecker product. Making the transformation $\underline{Y}_i = X\underline{\beta}_i$, it is seen that $\underline{Y}_i \sim CN_p(\underline{\emptyset}, \underline{\beta}_i'\underline{\beta}_i \otimes \Sigma)$ for $i = 1, 2, \cdots , r$ which is the same as $\underline{Y}_i \sim CN_p(\underline{\emptyset}, \Sigma)$ since $\underline{\beta}_i'\underline{\beta}_i = 1$ and the $\underline{Y}_1, \underline{Y}_2, \cdots , \underline{Y}_r$ form an independent set of $CN_p(\underline{0}, \Sigma)$ since $\underline{\beta}_i'\underline{\beta}_j = 0$, thus

$$XAX^{*\prime} = \sum_{i=1}^{r} X\underline{\beta}_i \underline{\beta}_i' X^{*\prime} = \sum_{i=1}^{r} \underline{Y}_i \underline{Y}_i^{*\prime} \sim CW_p(XAX^{*\prime}|r|\Sigma) \quad .$$

This result will be stated formally as the following theorem.

<u>Theorem 2.4</u>. $X = [\underline{\xi}_1, \underline{\xi}_2, \cdots , \underline{\xi}_n]$ where $\underline{\xi}_i$ are distributed as Independent $CN_p(\underline{0}, \Sigma)$, then for A a real indepotent matrix of rank $r \le p$, the matrix of Hermitian forms $XAX^{*\prime}$ is distributed as a $CW_p(XAX^{*\prime}|r|\Sigma_p)$.

2. Bartlett Decomposition of a Complex Wishart Matrix

Kshirsagar (1959) has presented an elegant method of deriving the Bartlett decomposition of a real Wishart matrix, i.e., expressing the Wishart matrix as the product of a lower triangular matrix times its transpose and giving the distribution of the elements of the lower triangular matrix by using random orthogonal transformations. This has been achieved in the case of a complex Wishart matrix by following the above procedure. The minor changes that are encountered are pointed out at the conclusion of this section.

Let $\underline{\xi}_1, \underline{\xi}_2, \cdots , \underline{\xi}_n$ where $\underline{\xi}_i' = [\xi_{1i}, \xi_{2i}, \cdots , \xi_{pi}]$, denote n independent and identically distributed p-variate complex normal vectors with probability density function given by $CN_p(\underset{\sim}{\emptyset}, I_p)$. Denote the matrix of the vectors by X, i.e., $X = [\underset{\sim}{\xi}_1, \underset{\sim}{\xi}_2, \cdots, \underset{\sim}{\xi}_n]_{p\times n}$.

Let A denote the matrix defined by

$$A = XX^{*'} = \sum_{i=1}^{n} \underset{\sim}{\xi}_i \underset{\sim}{\xi}_i^{*'}$$

then it has been shown [Khatri (1965)] that the matrix A is distributed as $CW_p(A|n|I_p)$.

Now consider the vectors defined by

$$\underset{\sim}{Z}_1 = \begin{bmatrix} \xi_{11} \\ \xi_{12} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{1n} \end{bmatrix} ; \quad \underset{\sim}{Z}_2 = \begin{bmatrix} \xi_{21} \\ \xi_{22} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{2n} \end{bmatrix} ; \cdots ; \quad \underset{\sim}{Z}_p = \begin{bmatrix} \xi_{p1} \\ \xi_{p2} \\ \cdot \\ \cdot \\ \cdot \\ \xi_{pn} \end{bmatrix} \qquad (2.2.0)$$

which are the vectors of observations on each of the p-variates. Each of the variables ξ_{ir} is independent and distributed as $CN_1(0, 1)$.

Orthogonalize these vectors, $\underset{\sim}{Z}_1, \cdots, \underset{\sim}{Z}_p$, by using the Gram-Schmitt process such that the new vectors form an orthonormal basis. The new vectors, $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \cdots, \underset{\sim}{Y}_p$, may be represented as

$$\begin{aligned} \underset{\sim}{Y}_1 &= b^{11}\underset{\sim}{Z}_1 \\ \underset{\sim}{Y}_2 &= b^{21}\underset{\sim}{Z}_1 + b^{22}\underset{\sim}{Z}_2 \\ &\cdot \\ &\cdot \\ &\cdot \\ \underset{\sim}{Y}_p &= b^{p1}\underset{\sim}{Z}_1 + b^{p2}\underset{\sim}{Z}_2 + \cdots + b^{pp}\underset{\sim}{Z}_p \quad . \end{aligned}$$

The b^{ii} for $i = 1, 2, \cdots, p$ can be chosen to be real while the b^{kj}, $k > j$, $k = 1, 2, \cdots, p$ $j = 1, 2, \cdots, p$ will in general be complex. To see this consider [Hohn (1958)]

$$\underline{Y}_1 = b^{11}\underline{Z}_1$$

and since b^{11} is to be chosen such that $\underline{Y}_1^{*'}\underline{Y}_1 = 1$, this implies $b^{11} = \frac{1}{\sqrt{Z_1^{*'}Z_1}}$ which is real. Now

$$\underline{Y}_2 = b^{21}\underline{Z}_1 + b^{22}\underline{Z}_2$$

and with the condition, $\underline{Y}_1^{*'}\underline{Y}_2 = 0$, this implies

$$b^{21}b^{11}\underline{Z}_1^{*'} \cdot \underline{Z}_1 + b^{11}b^{22} \cdot \underline{Z}_1^{*'}\underline{Z}_2 = 0 \ .$$

This is a homogeneous equation in two unknowns, b^{21} and b^{22} . A complete solution is given by

$$b^{21} = -(Z_1^{*'} \cdot Z_2)t$$
$$b^{22} = (Z_1^{*'} \cdot Z_1)t$$

where t is an arbitrary parameter. Continuing in this fashion, it can be shown that b^{33} is proportional to $(Z_1^{*'}Z_1)(Z_2^{*'}Z_2)$. So if t is chosen to be real, we have b^{22} real, also b^{33} real, and all other b^{ii} will be real. Attention should be called to b^{kj}, $k > j$. For a real t, this does not imply that b^{kj} will be real. Thus the b_{ii}, $i = 1, 2, \cdots, p$, will be restricted to be real.

Then the above system may be written as

$$\begin{bmatrix} \underline{Y}_1' \\ \underline{Y}_2' \\ \cdot \\ \cdot \\ \cdot \\ \underline{Y}_p' \end{bmatrix}_{p\times n} = \begin{bmatrix} b^{11} & 0 & \cdots & 0 \\ b^{21} & b^{22} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b^{p1} & 0 & \cdots & b^{pp} \end{bmatrix}_{p\times p} \begin{bmatrix} \underline{Z}_1' \\ \underline{Z}_2' \\ \cdot \\ \cdot \\ \cdot \\ \underline{Z}_p' \end{bmatrix}_{p\times n} . \qquad (2.2.1)$$

Denote (2.2.1) as

$$Y_{p\times n} = \tilde{B}_{p\times p} X_{p\times n} .$$

Since

$$X = [\xi_1, \xi_2, \cdots , \xi_n]_{p\times n} = \begin{bmatrix} \underline{Z}_1' \\ \underline{Z}_2' \\ \cdot \\ \cdot \\ \cdot \\ \underline{Z}_p' \end{bmatrix}_{p\times n} .$$

Let $B = [b_{ij}]$. Since the diagonal elements of $\tilde{B}$ are real, then the diagonal elements of $\tilde{B}^{-1} = B$ must necessarily be real. Hence, b_{ii} for $i = 1, 2, \cdots , p$ are real elements and the reciprocal relation may be written as

$$X_{p\times n} = B_{p\times p} Y_{p\times n} \qquad (2.2.2)$$

or as

$$\begin{bmatrix} \underline{Z}_1' \\ \underline{Z}_2' \\ \cdot \\ \cdot \\ \cdot \\ \underline{Z}_p' \end{bmatrix} = \begin{bmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ b_{p1} & b_{p2} & \cdots & b_{pp} \end{bmatrix} \begin{bmatrix} \underline{Y}_1' \\ \underline{Y}_2' \\ \cdot \\ \cdot \\ \cdot \\ \underline{Y}_p' \end{bmatrix}$$

which implies that

$$\underset{-}{Z}_k = b_{k1}\underset{-}{Y}_1 + b_{k2}\underset{-}{Y}_2 + \cdots + b_{kk}\underset{-}{Y}_k \quad , \tag{2.2.3}$$

i.e., the $\underset{-}{Z}_k$ can be expressed in terms of the unitary vectors $\underset{-}{Y}_k$, $k = 1, \cdots, p$.

From (2.2.2) it is seen that the matrix X can be represented as the matrix B times the semi-unitary matrix Y, then

$$XX^{*\prime} = BYY^{*\prime}B^{*\prime} = BB^{*\prime} \quad , \text{ since } \quad YY^{*\prime} = I_p \quad .$$

The distribution of $XX^{*\prime}$ is $CW_p(A|n|I_p)$ thus $BB^{*\prime} \sim CW_p(A|n|I_p)$ or that $A = BB^{*\prime}$. Thus a complex Wishart matrix can be expressed in terms of a lower triangular matrix, whose diagonal elements are real and the non-zero off diagonal elements are complex, times the transpose of its conjugate.

Also note that

$$\underset{-}{Y}_j^{*\prime}\underset{-}{Z}_k = b_{kj} \qquad \begin{matrix} k = 1, \cdots, p \\ j = 1, 2, \cdots, k \end{matrix} \tag{2.2.4}$$

this result follows from (2.2.3).

Now keep $\underset{-}{Z}_1, \underset{-}{Z}_2, \cdots, \underset{-}{Z}_{k-1}$ fixed. This automatically fixes $\underset{-}{Y}_1, \underset{-}{Y}_2, \cdots, \underset{-}{Y}_{k-1}, b_{11}, \cdots, b_{k-1,1}, b_{12}, \cdots, b_{k-1,k-1}$ by virtue of relation (2.2.4). Also recall that the ξ_{ir} are distributed independently as $CN_1(0, 1)$. Consider, from (2.2.4), the b_{kt}, $t = 1, 2, \cdots, k-1$, which may be written as

$$b_{k1} = \underline{Y}_1^{*\prime}\underline{Z}_k \tag{2.2.5}$$
$$b_{k2} = \underline{Y}_2^{*\prime}\underline{Z}_k$$
$$\vdots$$
$$b_{k\ k-1} = \underline{Y}_{k-1}^{*\prime}\underline{Z}_k$$

and by holding $\underline{Z}_1, \underline{Z}_2, \cdots, \underline{Z}_{k-1}$ fixed, this becomes a semi-unitary transformation from $\xi_{k1}, \xi_{k2}, \cdots, \xi_{kn}$ to $b_{k1}, b_{k2}, \cdots, b_{k\ k-1}$. Observing that

$$[\underbrace{\underline{Z}_1, \underline{Z}_2, \cdots, \underline{Z}_{k-1}}_{\text{fixed}} \mid \underbrace{\underline{Z}_k, \cdots, \underline{Z}_p}_{\text{random}}]$$

automatically requires that

$$\left[\begin{array}{cccc|ccc} \xi_{11} & \xi_{21} & \cdots & \xi_{k-1\ 1} & \xi_{k1} & \cdots & \xi_{p1} \\ \xi_{12} & \xi_{22} & \cdots & \xi_{k-1\ 2} & \xi_{k2} & \cdots & \xi_{p2} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \xi_{1n} & \xi_{2n} & \cdots & \xi_{k-1\ n} & \xi_{kn} & \cdots & \xi_{pn} \end{array}\right]$$

fixed random

Thus

$$b_{kj} = \underline{Y}_j^{*\prime}\underline{Z}_k \quad \text{for } j = 1, 2, \cdots, k-1$$

is distributed as a linear combination of the ξ_{kt} where $t = 1, 2, \cdots, n$ and the ξ_{kt} are distributed as independently $CN_1(0, 1)$. Thus

$$b_{kj} \sim CN_1(0, 1) \tag{2.2.6}$$

since $\text{var}(b_{kj}) = \underline{Y}_j^{*\prime}\underline{Y}_j = 1$, and $E(b_{jk}) = 0$. Furthermore, the b_{kt} forms an independent set for $t = 1, 2, \cdots, k-1$. Observe that for $t \neq t'$

$$\text{cov}(b_{kt}, b_{kt'}) = \text{cov}(Y_t^{*\prime}Z_k, Y_{t'}^{*\prime}Z_k) = Y_t^{*\prime}Y_{t'} = 0 .$$

Thus (2.2.6) is the conditional distribution of the b_{kj} with $\underline{Z}_1, \underline{Z}_2, \cdots, \underline{Z}_{k-1}$ being fixed, the $\underline{Z}_t$'s do not enter into the distribution, so (2.2.6) is also the unconditional distribution of the b_{kj}'s, for $j < k$ and the b_{kj}'s form a mutually independent set. Hence, the result

$$b_{k-1}, \cdots, b_{k\ k-1} \sim \text{Independent } CN_1(0, 1) . \tag{2.2.7}$$

As stated previously, the transformation (2.2.5) is a semi-unitary transformation from $\xi_{k1}, \cdots, \xi_{kn}$ to $b_{k1}, b_{k2}, \cdots, b_{k\ k-1}$. This transformation can be completed with $[n - (k-1)]$ new random variables in such a way as to maintain independence between $b_{k1}, b_{k2}, \cdots, b_{k\ k-1}$, and the $[n - (k-1)]$ random variables, so that the complete transformation is unitary. Since the transformation is of full rank and unitary, then

$$\xi_{k1}^*\xi_{k1} + \xi_{k2}^*\xi_{k2} + \cdots + \xi_{kn}^*\xi_{kn} =$$

$$b_{k1}^*b_{k1} + \cdots + b_{k\ k-1}^*b_{k\ k-1} + [n - (k-1)] \text{ random variables.}$$

Rewrite the above as

$$\underline{Z}_k^{*\prime}\underline{Z}_k = b_{k1}^*b_{k1} + \cdots + b_{k\ k-1}^*b_{k\ k-1} + [n - (k-1)] \text{ random variables} \tag{2.2.8}$$

and note the independence among the variables on the left once more. Furthermore, (2.2.8) is distributed as a $\frac{\chi^2_{(2n)}}{2}$, but by (2.2.3), it is seen that

$$\underset{\sim}{Z}_k^{*\prime}\underset{\sim}{Z}_k = b_{k1}^* b_{k1} + \cdots + b_{k\,k-1}^* b_{k\,k-1} + b_{kk}^2 \quad .$$

Also, by virtue of (2.1.0), it is seen

$$b_{k1}^* b_{k1} + \cdots + b_{k\,k-1}^* b_{k\,k-1} \sim \frac{\chi^2_{2(k-1)}}{2} \quad .$$

Thus by Fisher's Lemma, it is seen that

$$b_{kk}^2 = \underset{\sim}{Z}_k^{*\prime}\underset{\sim}{Z}_k - (b_{k1}^* b_{k1} + \cdots + b_{k\,k-1}^* b_{k\,k-1}) \sim \frac{\chi^2_{2[n - (k-1)]}}{2}$$

and b_{kk}^2 is independent of $b_{k1}, \cdots, b_{k\,k-1}$ and $\underset{\sim}{Z}_1, \underset{\sim}{Z}_2, \cdots, \underset{\sim}{Z}_{k-1}$. This result is true for every $k = 1, 2, \cdots, p$.

In summary, it is seen that a complex Wishart matrix A, may be represented as the product of a lower triangular matrix, $B = [b_{kj}]$, where b_{kk}, $k = 1, 2, \cdots p$, are real and b_{kj}, $k > j$, are complex times the transpose of its conjugate where the

$$b_{kj} \sim \text{Independent } CN_1(0, 1) \quad \text{for} \quad j < k,\ k = 1, 2, \cdots, p$$

and

$$b_{kk}^2 \sim \frac{\chi^2_{2[n - (k-1)]}}{2} \quad .$$

These results are similar to that of the decomposition of a real Wishart matrix in that the $b_{kj} \sim$ Independent $N(0, 1)$ for $j < k$, $k = 1, 2, \cdots, p$ and $b_{kk}^2 \sim \chi^2_{[n - (k-1)]}$.

3. Coherence Between Two Complex Random Variables

Ruben (1966) obtained, in the case of real bivariate normal situation, a very good approximation to the distribution of the correlation coefficient. This was done by expressing the sample correlation coefficient as a ratio of a linear combination of a standardized normal variate and a chi-variate to another chi-variate, where all three were mutually independent. This work has been extended to the multivariate situation by Kshirsagar (1969), who gives a matrix representation of this result for correlation between two vectors.

Following Ruben's work in the bivariate case and Kshirsagar's work in the multivariate case, similar results have been obtained in the complex case. These will be presented here.

Consider a random sample $\xi_1, \xi_2, \cdots, \xi_n$ of size n from a $CN_2(\underline{\mu}, \Sigma)$ where

$$\xi_k = \begin{bmatrix} z_{1k} \\ z_{2k} \end{bmatrix} ; \quad \underline{\mu} = \begin{bmatrix} \mu_1 + i\nu_1 \\ \mu_2 + i\nu_2 \end{bmatrix} ; \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} .$$

The Wishart matrix of sum of squares of products is given by

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \sim CW_2(A|n-1|\Sigma) ,$$

where

$$a_{ij} = \sum_{j=1}^{n} (z_{ij} - \bar{z}_i)(z_{ij} - \bar{z}_i)^*$$

and

$$\bar{z}_i = \frac{\sum_{k=1}^{n} z_{ik}}{n} \quad \text{for } i = 1, 2 .$$

The coherence between z_1 and z_2 is given by

$$\rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$$

and the sample coherence is given by

$$r = \frac{a_{12}}{\sqrt{a_{11}a_{22}}} \quad .$$

Define

$$\tilde{\rho} = \frac{\rho}{(1-\rho\rho^*)^{1/2}}$$

and

$$\tilde{r} = \frac{r}{(1-rr^*)^{1/2}} \quad . \qquad (2.3.0)$$

It is the quantity $\tilde{r}$ whose distribution will be approximated.

Now make the transformation from $\underline{\xi}$ to $\underline{Y}$ by the following transformation

$$\underline{Y}_{2\times 1} = C(\underline{\xi} - \underline{\mu})$$

where C is given by

$$C = \begin{bmatrix} \frac{1}{\sigma_{11}} & 0 \\ \frac{-\rho^*}{\sigma_{11}(1-\rho\rho^*)^{1/2}} & \frac{1}{\sigma_{22}(1-\rho\rho^*)^{1/2}} \end{bmatrix} \quad .$$

From Theorem 2.1, the distribution of $\underline{Y}$ is given by $CN_2(\underline{0}, C\Sigma C' = I_2)$ and, furthermore, the Wishart matrix A has been reduced to its canonical form, i.e., $D = CAC^{*\prime} \sim CW_2(D|n-1|I_2)$. These results are summarized in the following table

Variable	ξ	$\underline{Y} = C(\xi - \underline{\mu})$
mean	$\underline{\mu}$	$\underline{0}$
variance	Σ	I
Wishart matrix	A	$D = CAC^{*\prime}$

The Wishart matrix $D = CAC^{*\prime}$ has the form

$$D = \begin{bmatrix} \dfrac{a_{11}}{\sigma_{11}^2} & \dfrac{-\rho a_{11}}{\sigma_{11}^2(1-\rho\rho^*)^{1/2}} + \dfrac{a_{12}}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)^{1/2}} \\ \dfrac{-\rho^* a_{11}}{\sigma_{11}^2(1-\rho\rho^*)^{1/2}} + \dfrac{a_{12}^*}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)^{1/2}} & \delta \end{bmatrix}$$

where

$$\delta = \frac{a_{11}\rho\rho^*}{\sigma_{11}^2(1-\rho\rho^*)} - \frac{a_{12}^*\rho}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)} - \frac{a_{12}\rho^*}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)} + \frac{a_{22}}{\sigma_{22}^2(1-\rho\rho^*)} .$$

Since $D \sim CW_p(D|n-1|I_2)$, it may be written as $D = BB^{*\prime}$ where B is a lower triangular matrix given by the Bartlett decomposition. $BB^{*\prime}$ is given by (observing the b_{ii} for i = 1, 2 are real)

$$\begin{bmatrix} b_{11} & 0 \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{21}^* \\ 0 & b_{22} \end{bmatrix} = \begin{bmatrix} b_{11}^2 & b_{11}b_{21}^* \\ b_{21}b_{11}^* & (b_{21}b_{21}^* + b_{22}^2) \end{bmatrix}$$

and equating the elements of $BB^{*'}$ with the corresponding elements of $D = CAC^{*'}$, the following system of equations is obtained.

$$b_{11}^2 = \frac{a_{11}}{\sigma_{11}^2}$$

$$b_{11}b_{21}^* = \frac{a_{12}}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)^{1/2}} - \frac{\rho a_{11}}{\sigma_{11}^2(1-\rho\rho^*)^{1/2}}$$

$$b_{12}b_{12}^* + b_{22}^2 = \frac{a_{11}\rho\rho^*}{\sigma_{11}^2(1-\rho\rho^*)} - \frac{a_{12}^*\rho}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)} - \frac{a_{12}\rho^*}{\sigma_{11}\sigma_{22}(1-\rho\rho^*)} + \frac{a_{22}}{\sigma_{22}^2(1-\rho\rho^*)} .$$

Solving the above system for the a_{ij} $(i = 1, 2 ; j = 1, 2)$ the following solutions are obtained:

$$\begin{cases} a_{11} = \sigma_{11}^2 \, b_{11}^2 \\ a_{12} = \sigma_{11}\sigma_{22}(1-\rho\rho^*)^{1/2}(b_{11}\, b_{12} + \tilde{\rho}\, b_{11}^2) \\ a_{22} = \sigma_{22}^2(1-\rho\rho^*)\left[(b_{12}b_{12}^* + b_{22}^2) + \tilde{\rho}\, b_{11}b_{12}^* + \tilde{\rho}^* b_{11}b_{12} + \tilde{\rho}\,\tilde{\rho}^* b_{11}^2\right] . \end{cases} \quad (2.3.1)$$

Now, from Section 2, it has been shown that the

$$b_{11}^2 \sim \frac{\chi^2_{2(n-1)}}{2}$$

$$b_{22}^2 \sim \frac{\chi^2_{2(n-2)}}{2}$$

$$b_{12} \sim CN_1(0, 1)$$

and all are mutually independent. Upon substituting (2.3.1) into

(2.3.0), the following representation for $\tilde{r}$ is obtained

$$\tilde{r} = \frac{b_{12} + \tilde{\rho}\, b_{11}}{b_{22}}$$

which can be rewritten as

$$\tilde{r} = \frac{CN_1(0,\ 1) + \tilde{\rho}\, \frac{\chi_{2n-2}}{\sqrt{2}}}{\frac{\chi_{2n-4}}{\sqrt{2}}} \qquad (2.3.2)$$

where χ_v denotes a real chi-variable with v degrees of freedom. Attention should be called to the distribution of $\tilde{r}$ in the null case, i.e., $\rho = 0$,

$$\tilde{r} = \frac{CN(0,\ 1)}{\frac{\chi_{2n-4}}{\sqrt{2}}}$$

which has the appearance of what could be called a complex t.

Consider a probability statement about the complex random variable defined by (2.3.2)

$$P(\tilde{r} < \tilde{r}_0) = P\left(\frac{CN(0,\ 1) + \tilde{\rho}\, \frac{\chi_{2n-2}}{\sqrt{2}}}{\frac{\chi_{2n-4}}{\sqrt{2}}} < \tilde{r}_0\right)$$

or rewritten as

$$P(L < 0)$$

where

$$L = CN(0,\ 1) + \tilde{\rho}\, \frac{\chi_{2n-2}}{\sqrt{2}} - \tilde{\sigma}_0\, \frac{\chi_{2n-4}}{\sqrt{2}} \ . \qquad (2.3.3)$$

Using Fisher's normal approximation for a chi-variable with v degrees of freedom, i.e., $\chi \sim N\left(\sqrt{\frac{2v-1}{2}}, \frac{1}{2}\right)$, (2.3.3) can be rewritten as

$$L = CN(0, 1) + \frac{\tilde{\rho}}{\sqrt{2}} N\left(\sqrt{\frac{4n-3}{2}}, \frac{1}{2}\right) - \frac{\tilde{r}_0}{\sqrt{2}} N\left(\sqrt{\frac{4n-5}{2}}, \frac{1}{2}\right) .$$

Thus L is a linear combination of independent complex normal variables, which is also a complex normal random variable. The mean and variance of L is given by

$$E(L) = \tilde{\rho}\sqrt{\frac{4n-3}{4}} - \tilde{r}_0\sqrt{\frac{4n-5}{4}}$$

$$V(L) = 1 + \frac{1}{4}\tilde{\rho}\,\tilde{\rho}^* + \frac{1}{4}\tilde{r}_0\tilde{r}_0^*$$

and

$$P(L < 0) = \int_{-\infty}^{0} \frac{1}{\Pi\,[V(L)]} \exp\left[-\left(1 - E(L)\right)^*\left(L - E(L)\right)\left(V(L)\right)^{-1}_{dL}\right]$$

$$= \int_{-\infty}^{\frac{-E(L)}{V(L)}} \frac{1}{\Pi} e^{-uu^*}\, du$$

$$= \Phi_{CN}\left(\frac{-E(L)}{V(L)}\right) .$$

Replacing $\tilde{r}_0$ by $\tilde{r}$ in $-E(L)/V(L)$, then

$$\frac{\tilde{r}\sqrt{\frac{4n-5}{4}} - \tilde{\rho}\sqrt{\frac{4n-3}{4}}}{1 + \frac{1}{4}\tilde{\rho}\,\tilde{\rho}^* + \frac{1}{4}\tilde{r}\,\tilde{r}^*} \sim CN_1(0, 1) .$$

Now following Kshirsagar's (1969) work, a matrix representation for coherence between two complex vectors will be derived. Let the variance-covariance of two complex vectors $\underset{\sim}{X}_{p\times 1}$ and $\underset{\sim}{Z}_{q\times 1}$ be given by

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]_{(p+q)\times(p+q)} \begin{array}{l} p \\ q \end{array} \qquad \begin{array}{cc} p & q \end{array} \tag{2.3.4}$$

and let the matrix of corrected sum of squares (s.s.) and sum of products (s.p.) of observations from a sample of size n+1 on these variables be given by

$$S = \left[\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right]_{(p+q)\times(p+q)} \begin{array}{l} p \\ q \end{array} \qquad \begin{array}{cc} p & q \end{array} \tag{2.3.5}$$

which is based on n degrees of freedom. Then the following matrices can be obtained:

$\underset{\sim}{B}_0$ = matrix of regression coefficients $S_{12}S_{11}^{-1}$, of $\underset{\sim}{X}$ on $\underset{\sim}{Z}$;

B = matrix of s.s. and s.p., due to regression $S_{12}S_{22}^{-1}S_{21}$;

A = "residual" s.s. and s.p. matrix $S_{11} - S_{12}S_{22}^{-1}S_{21} = S_{11\cdot 2}$;

$A+B$ = "total" matrix S_{11} .

The corresponding matrices for the population are $\underset{\sim}{\beta} = \Sigma_{12}\Sigma_{22}^{-1}$, $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, and $\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$, respectively.

If the vector $\underline{Y}_{(p+q)\times 1}$ defined by $\underline{Y}' = [\underline{X}'|\underline{Z}']$ is distributed as $\ddot{CN}_{(p+q)}(\underline{\mu}, \Sigma)$ where $\underline{\mu}' = [\underline{\mu}_X|\underline{\mu}_Z]$ and Σ is defined by (2.3.4), then the matrix S, defined by (2.3.5) is distributed as $CW_{(p+q)}(S|n|\Sigma)$ and from this density it has been shown by Kabe (1966) that

$$S_{11\cdot 2} \sim CW_p\left(S_{11\cdot 2}|n-q|\Sigma_{11\cdot 2}\right) \qquad (2.3.6)$$

and that $S_{11\cdot 2}$ is distributed independently of B_0 and S_{22}. Also, Kabe (1966) has shown that

$$S_{22} \sim CW_q(S_{22}|n|\Sigma_{22}) \qquad (2.3.7)$$

and that $\underline{B}_0$, for fixed S_{22}, is distributed as

$$CN_p(\underline{\beta}, \Sigma_{11\cdot 2}) \qquad (2.3.8)$$

and is independent of $S_{11\cdot 2}$.

Now let Φ, η, M, F, C, and K be lower triangular matrices such that

$$\Sigma_{22} = \Phi\Phi^{*\prime}$$
$$\Sigma_{11\cdot 2} = \eta\eta^{*\prime}$$
$$S_{22} = MM^{*\prime}$$
$$A = FF^{*\prime}$$
$$B = KK^{*\prime}$$
$$A+B = CC^{*\prime} \quad .$$

Define further

$$U = \eta^{-1}(B_0 - B)M \qquad (2.3.9)$$

$$V = \Phi^{-1}M$$

$$W = \eta^{-1}F$$

$$\tilde{P} = \eta^{-1}\Sigma_{12}\Phi^{*\prime -1}$$

$$\tilde{R} = F^{-1}S_{12}M^{*\prime -1}$$

$$R = C^{-1}K$$

$$L = RR^{*\prime} = C^{-1}BC^{*\prime -1} \quad .$$

Transform to U, V, and W from B_0, Σ_{22}, and $S_{11\cdot 2}$, respectively, in (2.3.6), (2.3.7), and (2.3.8). Then it can be shown

(1) u_{ij} ($i = 1, \cdots, p$; $j = 1, 2, \cdots, q$), the pq variables in U are independent CN(0, 1);

(2) v_{jj} ($j = 1, \cdots, q$), the diagonal elements of the lower triangular matrix V are independent $\chi_{2(n-j+1)}/\sqrt{2}$ and the off diagonal element v_{kj} for $k > j$ are independent CN(0, 1) and independent v_{jj} ;

(3) w_{ii} ($i = 1, 2, \cdots, p$), the diagonal elements of the lower triangular matrix W are independent $\chi_{2(n-i+1)}/\sqrt{2}$ variates, while w_{ij}, $i > j$, are independent CN(0, 1), independent of w_{ii} also.

Referring back to the transformations in (2.3.9), it can be seen $L = RR^{*\prime}$ is the matrix generalization of r^2, which is the ratio of regression sum of squares to the total sum of squares and $\tilde{R}\tilde{R}^{*'}$ is the matrix generalization of $r^2/1-r^2$, the ratio of the regression sum of squares to the residual sum of squares. Observing that

$$U + \tilde{P}V = \eta^{-1}S_{12}M^{*'-1}$$

where $\tilde{P}$ is the population matrix corresponding to $\tilde{R}$, then

$$\tilde{R} = W^{-1}(U + \tilde{P}V) \qquad (2.3.10)$$

which is the matrix analogue to the result (2.3.2), i.e.,

$$\tilde{r} = \frac{CN_1(0,\ 1) + \tilde{\rho}\ \frac{\chi_{2n-2}}{\sqrt{2}}}{\frac{\chi_{2n-4}}{\sqrt{2}}}$$

which corresponds to Ruben's result in the real case. Then

$$\xi = U + \tilde{P}V - W\tilde{R}_0 \qquad (2.3.11)$$

is the matrix analogue of result (2.3.3), and also corresponds to Kshirsagar's result in the real case.

From this result (2.3.11), one can obtain results similar to Ruben's about every element of the matrix ξ, however in multivariate analysis, one prefers an overall criterion based on the whole matrix rather than individual results on the elements. Hence, no attempt is made to pursue this further; instead, we shall be considering overall criteria such as Wilks' Λ in the complex case.

4. Multiple Coherence

In the case of real multivariate normal analysis, an extensive amount of work has been done relating to the multiple correlation coefficient, i.e., the maximum correlation between a random variable x_p and a linear combination of a set of variables $x_1, x_2, \cdots, x_{p-1}$. If , where

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \hline \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{array}\right] = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \sigma_{pp} \end{array}\right]$$

denotes the variance-covariance structure of this set of variables $x_1, x_2, \cdots, x_p$, it can be shown that the square of the maximum correlation coefficient is given by

$$P^2_{x_p(x_1, \cdots, x_{p-1})} = \frac{\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}}{\sigma_{pp}}$$

and the same multiple correlation is given by

$$R^2_{x_p(x_1, \cdots, x_{p-1})} = \frac{A_{21}A_{11}^{-1}A_{12}}{a_{pp}}$$

where

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & a_{pp} \end{array}\right]$$

is the sample variance-covariance matrix based on a sample of size n. Hodgson (1967), following Ruben's (1966) approach, has shown $R^2/(1-R^2)$ is distributed as

$$\frac{\left[\chi^2_{k-1} + \left\{U + \left(P(1-P^2)\right)^{-\frac{1}{2}} \chi_{n-1}\right\}^2\right]}{\chi^2_{n-k-1}}$$

where the χ^2 and the unit normal variate U are mutually independent. From this, Hodgson has obtained a normal approximation for the statirtic $R^2/(1-R^2)$. These results have been achieved in a somewhat similar fashion for the multiple coherence in a multivariate complex normal situation.

Suppose $\underline{\xi} \sim CN_p(\underline{\mu}, \Sigma)$. Without loss of generality, it will be assumed that $\underline{\mu} = \underline{0}$ and Σ has the following special form

$$\Sigma = \begin{bmatrix} 1 & 0 & \cdots & 0 & P \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ P & 0 & \cdots & 0 & 1 \end{bmatrix}$$

where P is the true multiple coherence between ξ_p and $\xi_1, \xi_2, \cdots, \xi_{p-1}$, as defined in (1.2.6). This can be shown as below:

Let $\underline{\xi} \sim CN_p(\underline{\mu}, \Sigma)$, where

$$\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \cdot \\ \cdot \\ \cdot \\ \xi_p \end{bmatrix} \quad \text{and} \quad \Sigma = \left[\begin{array}{ccc|c} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \hline \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{array}\right] = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \sigma_{pp} \end{array}\right] .$$

Observing that Σ_{11} is an hermitian positive definite matrix, then there exists a D such that $\Sigma_{11} = DD^{*'}$ and $D^{-1}\Sigma_{11}D^{*'-1} = I_{p-1}$. Make the following transformation from $\underline{\xi} \to \underline{Y}$ by $\underline{Y} = A(\underline{\xi} - \underline{\mu})$, where A has the following form

$$A = \left[\begin{array}{c|c} D^{-1} & \underline{\emptyset} \\ \hline \underline{\emptyset}' & \dfrac{1}{\sigma_{pp}^{1/2}} \end{array}\right] .$$

then $E(\underline{Y}_{p\times 1}) = \underline{\emptyset}$ and $\mathrm{var}(\underline{Y}) = A\Sigma A^{*'}$, observe

$$A\Sigma A^{*'} = \left[\begin{array}{c|c} I_{p-1} & D^{-1}\Sigma_{12}\dfrac{1}{\sigma_{pp}^{1/2}} \\ \hline \Sigma_{12}' D^{*'-1}\dfrac{1}{\sigma_{pp}^{1/2}} & 1 \end{array}\right] .$$

Furthermore, the multiple coherence remains invariant under this transformation, since

$$P^2_{\xi_p(\xi_1, \xi_2, \cdots , \xi_{p-1})} = \frac{\Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}}{\sigma_{pp}}$$

and

$$\rho^2_{Y_p}(Y_1, Y_2, \cdots, Y_{p-1}) = \frac{\Sigma_{21} D^{*\prime -1} D^{-1} \Sigma_{12}}{\sigma_{pp}} = \frac{\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}}{\sigma_{pp}} ,$$

i.e.,

$$\rho^2_{\xi_p}(\xi_1, \cdots, \xi_{p-1}) = \rho^2_{Y_p}(Y_1, Y_2, \cdots, Y_{p-1}) .$$

Thus we have

$$\underline{Y}_{p\times 1} \sim CN_p(\underline{\emptyset}, \varepsilon)$$

where

$$\varepsilon = \left[\begin{array}{c|c} I_{p-1} & \begin{matrix} \varepsilon_{1p} \\ \vdots \\ \varepsilon_{p-1p} \end{matrix} \\ \hline \varepsilon_{p1} \cdots \varepsilon_{p-1p} & 1 \end{array}\right] .$$

Consider the true regression of Y_p on $Y_1, Y_2, \cdots, Y_{p-1}$, which can be written as

$$E(Y_p | Y_1, Y_2, \cdots, Y_{p-1}) = \beta_1 Y_1 + \beta_2 Y_2 + \cdots + \beta_{p-1} Y_{p-1} = \underline{\beta}' Y^{+}$$

where $\beta' = \varepsilon' I_{p-1} = [\varepsilon_{p1}, \varepsilon_{p2}, \cdots, \varepsilon_{p-1\,p}]$. Make the following transformation

$$x_1 = \frac{\beta'}{\sqrt{\beta'\beta^*}} \underline{Y}^+$$

$$x_2 = \frac{\underline{a}_2'}{\sqrt{\underline{a}_2'\underline{a}_2^*}} \underline{Y}^+$$

$$\vdots$$

$$x_{p-1} = \frac{\underline{a}_{p-1}'}{\sqrt{\underline{a}_{p-1}'\underline{a}_{p-1}^*}} \underline{Y}^+$$

$$x_p = Y_p$$

where $\underline{a}_2, \underline{a}_3, \cdots, \underline{a}_{p-1}$ are so chosen that

$$\underline{a}_2'\underline{\varepsilon} = \underline{a}_3'\underline{\varepsilon} = \cdots = \underline{a}_{p-1}'\underline{\varepsilon} = 0 \quad ,$$

i.e., find p-2 vector orthogonal to $\underline{\varepsilon}$, and noting the following relationships among the x's .

$$\text{var}(x_i) = 1 \quad \text{for} \quad i = 1, 2, \cdots , p$$

$$\text{cov}(x_i, x_j) = a_i^* a_j = 0 \quad \text{for} \quad i \neq j;\ i = 1, 2, \cdots , p-1;\ j = 1, 2, \cdots , p-1$$

$$\text{cov}(x_1, x_i) = \beta_j^{*\prime} a_i = 0 \quad \text{for} \quad i = 2, \cdots , p-1$$

$$\text{cov}(x_p, x_i) = \beta_j^{*\prime} a_i = 0 \quad \text{for} \quad i = 2, \cdots , p-1$$

and

$$\text{cov}(x_1, x_p) = \text{cov}\left(Y_p, \frac{1}{\sqrt{\beta^{*\prime}\beta}} \beta'\underline{Y}^+\right) = \frac{\beta^{*\prime}\beta}{\sqrt{\beta^{*\prime}\beta}}$$

$$= P \quad .$$

Thus we have

$$x \sim CN_p(\underline{0}, \ddagger)$$

where

$$\Sigma = \begin{bmatrix} 1 & 0 & \cdots & 0 & P \\ 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 0 \\ P & 0 & \cdots & 0 & 1 \end{bmatrix} \quad . \qquad (2.4.0)$$

Let $\underline{\xi}_1, \underline{\xi}_2, \cdots, \underline{\xi}_n$ denote a random sample of size n from a $CN_p(\underline{0}, \Sigma)$ where Σ is of the form given by (2.4.0). Let A_ξ denote the complex Wishart matrix associated with this sample, i.e.,

$$A_\xi = \sum_{i=1}^{n} \underline{\xi}_i \, \underline{\xi}_i^{*\prime} \sim CW_p(A|n|\Sigma) \quad .$$

It is desired to look at the Bartlett decomposition of this Wishart matrix, but the distribution of A_ξ is not in the canonical form. By making a transformation of the original sample, it is possible to express A_ξ as

$$A_\xi = CTT^{*\prime}C^{*\prime}$$

where T is a p x p triangular matrix whose elements are independent random variables, the off diagonal elements being $CN_1(0, 1)$ and the diagonal elements being χ variables. To see this, consider the following argument.

Let $\underline{\xi} \sim CN_p(\underline{0}, \Sigma)$, Σ is hermitian positive definite, thus $\Sigma = CC^{*\prime}$. Transform from $A_{\underline{\xi}}$ to $A_{\underline{Y}}$ by

$$A_{\underline{Y}} = C^{-1}A_{\underline{\xi}}C^{*\prime -1} \quad ,$$

where

$$\Sigma = CC^{*'}$$

or

$$A_{\underset{\sim}{\xi}} = CA_{\underset{\sim}{Y}}C^{*'} = CTT^{*'}C^{*'} \tag{2.4.1}$$

which is the desired result.

Now if $\underset{\sim}{\xi}_i \sim CN_p(\underset{\sim}{0}, \Sigma)$ and Σ is of the form (2.4.0), the matrix C, such that $C^{-1}\Sigma C^{*'-1} = I_p$ is given by

$$C = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 0 \\ P & 0 & \cdots & 0 & (1-P^2)^{1/2} \end{bmatrix}$$

and from (2.4.1), A is given by $A = CTT^{*'}C^{*'}$. First examine CT,

$$CT = \left[\begin{array}{c|c} I_{p-1} & \underset{\sim}{0} \\ \hline P\ 0 \cdots 0 & (1-P^2)^{1/2} \end{array}\right] \cdot \left[\begin{array}{cccc|c} t_{11} & 0 & \cdots & 0 & \\ t_{21} & t_{22} & & & \underset{\sim}{0} \\ t_{p-1\,1} & t_{p-1\,2} & \cdots & t_{p-1\,p-1} & \\ \hline t_{p1} & t_{p2} & \cdots & t_{p\,p-1} & t_{pp} \end{array}\right]$$

$$= \left[\begin{array}{c|c} I_{p-1} & \underset{\sim}{0} \\ \hline \underset{\sim}{\S}' & (1-P^2)^{1/2} \end{array}\right] \cdot \left[\begin{array}{c|c} T & \underset{\sim}{0}' \\ \hline \underset{\sim}{t}' & t_{pp} \end{array}\right]$$

$$= \left[\begin{array}{c|c} T & \underset{\sim}{0} \\ \hline \underset{\sim}{\S}'T + (1-P^2)^{1/2}\underset{\sim}{t}' & (1-P^2)^{1/2}t_{pp} \end{array}\right]$$

observing

$$\underline{\S}'T + (1-P^2)^{\frac{1}{2}}\underline{t}' = [pt_{11}, 0\ 0 \cdots 0] + \left[(1-P^2)^{\frac{1}{2}}t_{p1}, (1-P^2)^{\frac{1}{2}}t_{p2}, \cdots, (1-P^2)^{\frac{1}{2}}t_{pp}\right]$$

and letting X' represent this quantity, then

$$CT = \left[\begin{array}{c|c} T & \underline{0} \\ \hline X' & (1-P^2)^{1/2}t_{pp} \end{array}\right]$$

and CTT*'C*' is given as

$$CTT^{*\prime}C^{*\prime} = \left[\begin{array}{c|c} TT^{*\prime} & T\underline{X}^{*} \\ \hline \underline{X}'T^{*\prime} & (1-P^2)t_{pp}^2 \\ & + \underline{X}'\underline{X}^{*} \end{array}\right]$$

observing that

$$\underline{X}'\underline{X}^{*} = [Pt_{11} + (1-P^2)^{1/2}t'_{p1}][Pt_{11} + (1-P^2)^{1/2}t^{*}_{p1}]$$

$$+ (1-P^2)\sum_{i=2}^{p-1} t_{pi}t^{*}_{pi} \quad .$$

Hence

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & a_{pp} \end{array}\right] = \left[\begin{array}{c|c} TT^{*\prime} & TX^{*} \\ \hline X'T^{*\prime} & (1-P^2)t_{pp}^2 \\ & + X'X^{*} \end{array}\right] \quad . \qquad (2.4.2)$$

Interest lies with the quantity $R^2/(1-R^2)$. Observe that

$$1-R^2 = \frac{|A|}{|A_{11}| \cdot a_{pp}} \tag{2.4.3}$$

which may be rewritten

$$1-R^2 = \frac{(1-P^2) \prod_{i=1}^{p} t_{ii}^2}{|TT^{*'}| \cdot \{(1-P^2)t_{pp}^2 + \underline{X}'\underline{X}^*\}}$$

$$= \frac{(1-P^2)t_{pp}^2}{\{(1-P^2)t_{pp}^2 + X'X^*\}}$$

then

$$\frac{R^2}{1-R^2} = \frac{X'X^*}{(1-P^2)t_{pp}^2}$$

$$= \frac{|\{Pt_{11} + (1-P^2)^{1/2}t_p\}|^2 + (1-P^2)\sum_{i=2}^{p-1} t_{pi}t_{pi}^*}{(1-P^2)t_{pp}^2}$$

or

$$\frac{R^2}{1-R^2} = \frac{|t_{p1} + (P/(1-P^2)^{1/2})t_{11}|^2 + \sum_{i=2}^{p-1} t_{pi}t_{pi}^*}{t_{pp}^2} \tag{2.4.4}$$

where $|t_{p1} + (P/(1-P^2)^{1/2})t_{11}|^2$ represents the modulus of the quantity.

Let $\delta = \left(\frac{(1-P^2)^{1/2}}{P} t_{p1} + t_{11}\right)$ and using the fact that $t_{p1} \sim CN_1(0, 1)$ and t_{11} which is a chi-variable and can be approximated by a real $N\left(\sqrt{\frac{2(2n-2)-1}{4}}, \frac{1}{4}\right)$ the δ can be approximated by

$$\delta \sim CN_1\left(\sqrt{\frac{4n-5}{4}}\;,\; \frac{1}{4} + \frac{P^2}{(1-P^2)}\right)$$

then

$$\frac{R^2}{1-R^2} = \frac{\left(1 + \frac{1}{4}\lambda\right)\chi^2_{(2,\ \lambda)} + \chi^2_2 \sum\limits_{i=1}^{p-1} [n-(i-1)]}{\chi^2_2(n-p-1)}$$

where the noncentrality parameter λ is given by

$$\lambda = \frac{P^2}{(1-P^2)}$$

if $P = 0$, i.e., the true multiple coherence coefficient is zero, then

$$\frac{R^2}{1-R^2} = F\left[2 \sum_{k=1}^{p-1} \left\{[n-(k-1)],\ 2[n-(p-1)]\right\}\right] .$$

This is exact and follows from (2.4.4).

CHAPTER III

1. A Test for Equality of Means

In the case of k real p-variate normal populations, tests have been developed for testing hypotheses related to the means of these k populations. Test statistics have been developed by Roy, Lawley, Hotelling, Pillai, and Wilks. Wilks' statistic, usually referred to as Wilks' Λ, is the simplest one and is related to the likelihood ratio criterion. Thus it has the desirable properties associated with the likelihood ratio procedure. It is for these reasons that the complex analogue of Wilks' statistic will be used to develop tests for distinguishing between the means of k complex p-variate normal populations.

Let samples of size $n_1, n_2, \cdots, n_k$ be available from $k = q+1$ complex p-variate normal populations with means $\underline{\mu}(1), \underline{\mu}(2), \cdots, \underline{\mu}(k)$ and the same variance covariance structure given by Σ. Let $\underline{\xi}_{\alpha\ell}$ denote the vector of responses for the ℓ^{th} member of the α population ($\ell = 1, 2, \cdots, n_\alpha$; $\alpha = 1, 2, \cdots, k$) on which the observations are made. It is desired to develop a test for the hypothesis of equal means, i.e.,

$$H_0: \underline{\mu}(1) = \underline{\mu}(2) = \cdots = \underline{\mu}(k)$$

against the alternate hypothesis that the populations have different means. Represent the alternate hypothesis by H_a.

Giri (1965) has shown that the likelihood ratio criterion for certain hypotheses about complex multivariate normal populations possesses optimum properties which are characteristic of this test in the real case. These properties will not be investigated here, but since the likelihood ratio test has been found satisfactory for the problem in the real case, it will be used here in the complex case. The author feels that the desirable properties are maintained.

Letting $N = n_1 + n_2 + \cdots + n_k$, then the likelihood function, Q, may be written as

$$Q = \frac{1}{(\pi)^{NP}(|\Sigma|^N)} \exp\left\{- \sum_{\alpha=1}^{k} \sum_{\ell=1}^{n_\alpha} [\underline{\xi}_{\alpha\ell} - \underline{\mu}(\alpha)]^{*\prime}\Sigma^{-1}[\underline{\xi}_{\alpha\ell} - \underline{\mu}(\alpha)]\right\} .$$

The likelihood ratio statistic is given by

$${}_c\Lambda = \frac{\text{Max } Q_0}{\text{Max } Q}$$

where Max Q_0 is the maximum of the likelihood function under the assumption that the null hypothesis, H_0, is true and Max Q is the maximum value of the likelihood function over the entire parameter space, denoted by Ω. The test is given by rejecting H_0 when ${}_c\Lambda$ is less than some specified constant depending on the size of the test.

Using Giri's (1965) results, the Max Q_0 is obtained when

$$\hat{\underline{\mu}} = \frac{1}{N} \sum_{\alpha=1}^{k} \sum_{\ell=1}^{n_\alpha} \underline{\xi}_{\alpha\ell} = \underline{\xi}_{..}$$

and

$$\hat{\Sigma} = \frac{1}{N} T \quad \text{where} \quad T = \sum_{\alpha=1}^{k} \sum_{\ell=1}^{n_\alpha} (\underline{\xi}_{\alpha\ell} - \underline{\xi}_{..})(\underline{\xi}_{\alpha\ell} - \underline{\xi}_{..})^{*\prime} .$$

The Max Q is given when

$$\hat{\mu} = \frac{1}{n_\alpha} \sum_{\ell=1}^{n_\alpha} \underset{\sim}{\xi}_{\alpha\ell} = \underset{\sim}{\xi}_{\alpha\cdot}$$

and

$$\hat{\Sigma} = A = \sum_{\alpha=1}^{k} A_\alpha \quad \text{where} \quad A_\alpha = \sum_{\ell=1}^{n_\alpha} (\underset{\sim}{\xi}_{\alpha\ell} - \underset{\sim}{\xi}_{\alpha\cdot})(\underset{\sim}{\xi}_{\alpha\ell} - \underset{\sim}{\xi}_{\alpha\cdot})^{*'} \quad .$$

After some simplification, it can be shown that

$${}_c\Lambda = \frac{|A|}{|T|}$$

and observing that

$$T = A + B$$

where B is defined by

$$B = \sum_{\alpha=1}^{k} n_\alpha (\underset{\sim}{\xi}_{\alpha\cdot} - \underset{\sim}{\xi}_{\cdot\cdot})(\underset{\sim}{\xi}_{\alpha\cdot} - \underset{\sim}{\xi}_{\cdot\cdot})^{*'} \quad .$$

Then the likelihood ratio statistic is given by

$${}_c\Lambda = \frac{|A|}{|A+B|} \quad .$$

The statistic ${}_c\Lambda$ is the complex analogue of the result obtained by Wilks; it will be referred to as Wilks' Λ in the complex case and denoted as ${}_c\Lambda$.

Gupta (1971) has derived the exact distribution of ${}_c\Lambda$ when the alternate hypothesis is of unit rank. He gives the explicit expression for the case of p = 2, 3 and general degrees of freedom for the hypothesis and error.

The results obtained by application of the likelihood ratio principle may be summarized in the following complex multivariate analysis of variance table (CMANOVA).

Source of Variation	d.f.	s.s. and s.p.
Between groups	q	$B = \sum_{\alpha=1}^{k} n_\alpha (\underline{\xi}_{\alpha\cdot} - \underline{\xi}_{\cdot\cdot})(\underline{\xi}_{\alpha\cdot} - \underline{\xi}_{\cdot\cdot})^{*\prime}$
Within groups	n−q	A = †
TOTAL	n = N−1	$A+B = \sum_{\alpha=1}^{k} \sum_{\ell=1}^{n_\alpha} (\underline{\xi}_{\alpha\ell} - \underline{\xi}_{\cdot\cdot})(\underline{\xi}_{\alpha\ell} - \underline{\xi}_{\cdot\cdot})^{*\prime}$

†(by subtraction).

That the degrees of freedom are indeed q for the between groups and n−q for the within groups will now be shown by examining the distributions of the matrices B and A.

Define a p × N matrix X by

$$X = [\underline{\xi}_{11}, \underline{\xi}_{12}, \cdots, \underline{\xi}_{1n_1} | \underline{\xi}_{21}, \underline{\xi}_{22}, \cdots, \underline{\xi}_{2n_2} | \cdots | \underline{\xi}_{k1}, \underline{\xi}_{k2}, \cdots, \underline{\xi}_{kn_k}] .$$

Let E_{NN} denote an N × N matrix composed of all ones and let $\underline{e}_{N\times 1}$ denote a vector with all elements being unity. Also define $\underline{e}_\alpha$, an N × 1 vector, with zero everywhere except in the n_α positions; there are ones in all of the n_α positions, i.e.,

$$\underline{e}_\alpha' = [0\ 0 \cdots 0 \cdots \underbrace{1\ 1 \cdots 1}_{n_\alpha \text{ ones}} \cdots 0\ 0 \cdots 0]_{1\times N} .$$

Then the matrices A, B, and A+B may be written as

$$B = X\left[\sum_{\alpha=1}^{k} \frac{1}{n_\alpha} \underline{e}_\alpha \underline{e}_\alpha' - \frac{1}{N} E_{NN}\right]X^{*'} = XCX^{*'} \;; \quad (3.1.0)$$

$$A = X\left[I_N - \sum_{\alpha=1}^{k} \frac{1}{n_\alpha} \underline{e}_\alpha \underline{e}_\alpha'\right]X^{*'} = XDX^{*'} \;; \quad (3.1.1)$$

$$A+B = X\left[I_N - \frac{1}{N} E_{NN}\right]X^{*'} = XEX^{*'} \;. \quad (3.1.2)$$

It should be noted that the matrices C, D, and E are idempotent of ranks q, n-q, and n, respectively.

Observing that

$$E(X) = M = [\underline{\mu}(1), \underline{\mu}(1), \cdots, \underline{\mu}(1) | \cdots | \underline{\mu}(k), \underline{\mu}(k), \cdots, \underline{\mu}(k)] \quad (3.1.3)$$

and recalling that

$$\underline{\xi}_{\alpha\ell} \sim CN_p(\underline{\mu}(\alpha), \Sigma)$$

then by Theorem 2.4,

$$(X-M)C(X-M)^{*'} \sim CW_p\big((X-M)C(X-M)^{*'} | q | \Sigma\big) \;. \quad (3.1.4)$$

Thus the degrees of freedom associated with the between groups sum of squares (s.s.) and sum of products (s.p.) matrix B, is given by the rank of the idempotent matrix C which is q.

From (3.1.3), it can be shown that

$$E(XCX^{*'}) = q\Sigma + MCM^{*'}$$

and under the assumption of H_o, i.e., equal means for the k-complex p-variate populations, M can be written as

$$M = \underline{\mu}\underline{e}_{1\times N}$$

where $\underline{\mu}$ represents the common value. Observe that under H_0,

$$MC = \emptyset \ .$$

Then

$$\begin{aligned}(X-M)C(X-M)^{*'} &= (X-M)C(X^{*'}-M^{*'}) \\ &= XCX^{*'} - XCM^{*'} - MCX^{*'} + MCM^{*'} \\ &= XCX^{*'} \ .\end{aligned}$$

Thus

$$XCX^{*'} \sim CW_p(XCX^{*'}|q|\Sigma) \quad \text{if} \quad \underline{\mu}(1) = \underline{\mu}(2) = \cdots = \underline{\mu}(k) \ .$$

Now the distribution of $A = XDX^{*'}$ needs to be established. From Theorem 2.4, it is seen that

$$(X-M)D(X-M)^{*'} \sim CW_p\left((X-M)C(X-M)^{*'}|n-q|\Sigma\right) \ .$$

And the degrees of freedom associated with A are given by n-q, the rank of the idempotent matrix D.

Now

$$(X-M)D(X-M)^{*'} = XDX^{*'} - XDM^{*'} - MDX^{*'} + MDM^{*'}$$

and it can be shown that

$$MD = \emptyset \ .$$

Hence

$$(X-M)D(X-M)^{*'} = XDX^{*'} \sim CW_p(XDX^{*'}|n-q|\Sigma) \quad .$$

Attention should be called to the fact that the distribution of $A = XDX^{*'}$ is independent of the null hypothesis, i.e., regardless of whether H_o is true or not, the distribution of A is still $CW_p(A|n-q|\Sigma)$.

Also, it should be noted that A and B are distributed independently. Since, as in the real case, two forms $XCX^{*'}$ and $XDX^{*'}$ are independent if $CD = \emptyset$.

To summarize these results, it is seen that

$$B \sim CW_p(B|q|\Sigma) \quad \text{if} \quad \underline{\mu}(1) = \underline{\mu}(2) = \cdots = \underline{\mu}(k)$$

$$A \sim CW_p(A|n-q|\Sigma) \quad \text{independent of } H_o$$

and A is distributed independently of B. The test statistic, ${}_c\Lambda$, is given by the ratio of the determinates of the complex Wishart matrices A and A+B, i.e.,

$${}_c\Lambda = \frac{|A|}{|A+B|} \quad .$$

Khatri (1965) has shown that the distribution of ${}_c\Lambda$, where ${}_c\Lambda$ is defined as above, is the same as the product of p independent real beta variables, i.e.,

$${}_c\Lambda = \frac{|A|}{|A+B|} = \prod_{i=1}^{p} u_i$$

where $u_i \sim B(n-q - i+1, q)$ and all are independent for $i = 1, 2, \cdots , p$.

Furthermore, he has shown

$$-m \log {}_c\Lambda \stackrel{\cdot}{\sim} \chi^2_{2pq}$$

where m = (2n-q-p) and "$\stackrel{\cdot}{\sim}$" denotes "approximately distributed."

Hence if H_0 is true, $-m \log {}_c\Lambda$ will be distributed as an approximate chi-square variable with degrees of freedom given by 2pq.

To summarize the procedure for testing the hypothesis of equal means among k = q+1 complex p-variate normal populations, one must first calculate the matrices A and A+B. Then, consider the statistic ${}_c\Lambda$ given by

$${}_c\Lambda = \frac{|A|}{|A+B|}$$

and compute $-m \log {}_c\Lambda$. If $-m \log {}_c\Lambda$ is greater than χ^2_{2pq} at the desired α-level, then reject H_0, otherwise do not reject H_0 .

2. A Test for the Dimensionality of the Mean Space

In the last section a test for equality of the means from k p-variate complex normal populations with the same variance covariance structure was developed. The hypothesis, H_0, implies that the mean vectors of the k populations would lie in a complex space of zero dimension, i.e., represented by a point in the complex space. If H_0 is rejected in favor of H_a, then the means of the k populations may lie in a complex 1 dimensional space, that is they are collinear, or in a complex 2 dimensional space, that is they are coplanar, and so forth up to a complex k dimensional space.

In real multivariate analysis, it is important to know the dimensionality of the space spanned by the mean vectors for this is equivalent to knowing the number of linear discriminant functions that are necessary to distinguish among the k-populations. This is true also in the complex normal situation. In this section a test for the dimensionality of the complex space spanned by the mean vectors will be developed and will be shown to be equivalent to the number of linear discriminant functions needed to discriminate among the k populations.

Recalling the complex multivariate analysis of variance table that was developed in the previous section, and adding a column for the expectation of the s.s. and s.p. matrices, then the table is given by

Source of Variation	d.f.	s.s. and s.p.	E(s.s. and s.p.)
Between groups	q	$B = XCX^{*'}$	$E(B) = q\Sigma + MCM^{*'}$
Within groups	n-q	$A = XDX^{*'}$	$E(A) = (n-q)\Sigma$
TOTAL	n	A+B	

Now M was defined by (3.1.4) as

$$E(X) = M = [\underline{\mu}(1) \cdots \underline{\mu}(1) | \cdots | \underline{\mu}(k) \cdots \underline{\mu}(k)]_{p\times N} \ .$$

Denote the rank of M by r(M) and observe that

$$r(M) = r\big([\underline{\mu}(1), \underline{\mu}(2), \cdots, \underline{\mu}(k)]\big) = r(\mu)$$

which is equivalent to the number of linearly independent points representing the means of the k populations. Denote the k populations by $\pi_1, \pi_2, \cdots, \pi_k$ and consider a linear combination of $\underline{\xi}$, an observation from one of these populations, given by $\underline{\ell}^{*'}\underline{\xi}$. Suppose further that

$$\underline{\ell}^{*'}\underline{\mu}(1) = \underline{\ell}^{*'}\underline{\mu}(2) = \cdots = \underline{\ell}^{*'}\underline{\mu}(k) = \delta \ ,$$

i.e., that $\underset{\sim}{\ell}^{*'}\underset{\sim}{\xi}$ has the same mean in $\pi_1, \pi_2, \cdots, \pi_k$. Then this linear function given by $\underset{\sim}{\ell}^{*'}$ would not be useful in discriminating among the populations, i.e., one could not determine which population $\underset{\sim}{\xi}$ came from by examing $\underset{\sim}{\ell}^{*'}\underset{\sim}{\xi}$. Furthermore, observe that

$$\underset{\sim}{\ell}^{*'}M = [\delta\ \delta\ \cdots\ \delta] = \delta \cdot e_{1\times N}$$

then

$$\underset{\sim}{\ell}^{*'}MC = \delta\ e_{1\times N}\ C = \delta\ e_{1\times N}\left(\sum_{\alpha=1}^{k} \frac{1}{n_\alpha} \underset{\sim}{e}_\alpha \underset{\sim}{e}'_\alpha - \frac{1}{N} E_{NN}\right)$$

$$= \delta\ \underset{\sim}{0}' = \underset{\sim}{0}'$$

and hence

$$\underset{\sim}{\ell}^{*'}MCM^{*'}\underset{\sim}{\ell} = 0 \quad .$$

Conversely, if $\underset{\sim}{\ell}^{*'}MCM^{*'}\underset{\sim}{\ell} = 0$, this implies

$$\sum_{\alpha=1}^{k} n_\alpha (u_\alpha - \bar{u})(u_\alpha - \bar{u})^{*'} = 0 \tag{3.2.0}$$

where $\underset{\sim}{\ell}^{*'}\underset{\sim}{\mu}(\alpha) = u_\alpha$ and $\bar{u} = \sum_{\alpha=1}^{k} \frac{n_\alpha \underset{\sim}{\ell}^{*'}\underset{\sim}{\mu}(\alpha)}{\Sigma n_\alpha}$. But (3.2.0) implies that each term of the summation must be zero or that

$$u_\alpha = \bar{u} \quad \text{for all} \quad \alpha = {}^-, 2, \cdots, k \ .$$

From this it is seen that

$$\underset{\sim}{\ell}^{*'}\underset{\sim}{\mu}(1) = \underset{\sim}{\ell}^{*'}\underset{\sim}{\mu}(2) = \cdots = \underset{\sim}{\ell}^{*'}\underset{\sim}{\mu}(k) \quad .$$

So if $\underline{\ell}^{*\prime}\underline{\xi}$ has the same mean value in $\pi_1, \pi_2, \cdots, \pi_k$ then $\underline{\ell}^{*\prime}MC = 0$ and conversely if $\underline{\ell}^{*\prime}MC = 0$, then $\underline{\ell}^{*\prime}\underline{\xi}$ will have the same mean value in $\pi_1, \pi_2, \cdots, \pi_k$. In either case $\underline{\ell}^{*\prime}\underline{\xi}$ will not be useful as a discriminator.

Thus a linear function of $\underline{\xi}$, given by $\underline{\ell}^{*\prime}\underline{\xi}$, will not be a good discriminator if

$$\underline{\ell}^{*\prime}MC = 0 \;. \tag{3.2.1}$$

An adequate number of linear discriminators will be given by p minus the number of independent $\underline{\ell}$'s which satisfy (3.2.1). For any $\underline{\ell}$, such that (3.2.1) is satisfied, $\underline{\ell}^{*\prime}\underline{\xi}$ will be called a covariate. Observe that (3.2.1) is equivalent to

$$MCM^{*\prime}\underline{\ell} = 0 \;.$$

So the maximum number of such covariates is equal to the multiplicity of the zero characteristic roots of the matrix $MCM^{*\prime}$. Denote this multiplicity by s. Now the rank of $MCM^{*\prime}$ is equivalent to the number of non-zero characteristic roots of $MCM^{*\prime}$, say r, which is the number of linear discriminate functions needed. Hence, we have the number of discriminate functions needed given by

$$r = p-s \;.$$

Let the s covariates be given by

$$\underline{\ell}^{*\prime}(1)\underline{\xi},\ \underline{\ell}^{*\prime}(2)\underline{\xi},\ \cdots,\ \underline{\ell}^{*\prime}(s)\underline{\xi}$$

and let the matrix L be defined by

$$L = \begin{bmatrix} \underline{\ell}^{*\prime}(1) \\ \underline{\ell}^{*\prime}(2) \\ \cdot \\ \cdot \\ \cdot \\ \underline{\ell}^{*\prime}(s) \end{bmatrix}_{s\times p} \qquad (3.2.2)$$

then

$$L_{s\times p}\mu_{p\times k} = \underline{\nu}E_{1k} = [\underline{\nu}, \underline{\nu}, \cdots , \underline{\nu}]_{s\times k}$$

since $\underline{\ell}^{*\prime}(t)\underline{\xi}$ $(t = 1, 2, \cdots , s)$ is a covariate and this implies that $\underline{\ell}^{*\prime}(t)\underline{\mu}(\alpha) = \nu$, $(t = 1, 2, \cdots , s;\ \alpha = 1, 2, \cdots , k)$. Since $\underline{\ell}(1)$, $\underline{\ell}(2), \cdots , \underline{\ell}(s)$ are independent, this implies that the rank of L is s, i.e., $r(L) = s$. Thus the hypothesis of needing r discriminate functions is equivalent to the hypothesis that there exists a matrix L of rank s such that

$$L\mu = \underline{\nu}E_{1k}$$

for some unspecified $\underline{\nu}$. A test for this hypothesis will now be developed by using the likelihood ratio criterion.

Recall that the likelihood function, Q, was given as

$$Q = \frac{1}{\Pi^{PN}|\Sigma|^{N}} \exp\left(-\sum_{\alpha=1}^{k}\sum_{\ell=1}^{n_\alpha}\left(\underline{\xi}_{\alpha\ell} - \underline{\mu}(\alpha)\right)^{*\prime}\Sigma^{-1}\left(\underline{\xi}_{\alpha\ell} - \underline{\mu}(\alpha)\right)\right).$$

By adding and subtracting $\bar{\underline{\xi}}(\alpha) = \frac{1}{n_\alpha}\sum_{\ell=1}^{n_\alpha} \underline{\xi}_{\alpha\ell}$ and simplifying, Q may be written as

$$Q = \frac{1}{\Pi^{PN}|\Sigma|^N} \exp\left(-\sum_{\alpha=1}^{k}\sum_{\ell=1}^{n_\alpha}\left(\underline{\xi}_{\alpha\ell}-\bar{\underline{\xi}}(\alpha)\right)^{*\prime}\Sigma^{-1}\left(\underline{\xi}_{\alpha\ell}-\bar{\underline{\xi}}(\alpha)\right)\right)$$

$$\cdot \exp\left(-\sum_{\alpha=1}^{k} n_\alpha\left(\bar{\underline{\xi}}(\alpha)-\underline{\mu}(\alpha)\right)^{*\prime}\Sigma^{-1}\left(\bar{\underline{\xi}}(\alpha)-\underline{\mu}(\alpha)\right)\right) .$$

Assuming Σ to be known, the Q can be written as

$$Q = \left(\text{constant w.r.t. unknown parameters, } \underline{\mu}(\alpha)\right)$$

$$\cdot \exp\left(-\sum_{\alpha=1}^{k} n_\alpha\left(\bar{\underline{\xi}}(\alpha)-\underline{\mu}(\alpha)\right)^{*\prime}\Sigma^{-1}\left(\bar{\underline{\xi}}(\alpha)-\underline{\mu}(\alpha)\right)\right) .$$

The parameter space Ω consists of the 2pk parameters given by $\underline{\mu}(1)$, $\underline{\mu}(2), \cdots, \underline{\mu}(k)$. The maximum of Q with respect to Ω is given when $\hat{\underline{\mu}}(\alpha) = \bar{\underline{\xi}}(\alpha)$.

Now the likelihood function must be maximized subject to the conditions of the hypothesis of needing r discriminant functions. This is equivalent to the hypothesis that there exists a matrix L such that

$$H:\quad L\underline{\mu}(\alpha) = \underline{\nu} \qquad \alpha = 1, 2, \cdots, k$$

or

$$H:\quad L\mu = \underline{\nu}E_{1k}$$

which is equivalent to saying the $r(MCM^{*\prime}) = r$. In all cases it should be noted that the matrix L and vector $\underline{\nu}$ are not specified. The maximization under the assumption that H is true will be carried out through the following procedure.

Make the following transformation

$$L\bar{\underline{\xi}}(\alpha) = \underline{y}(\alpha) \qquad \alpha = 1, 2, \cdots, k$$

$$L_1\bar{\underline{\xi}}(\alpha) = \underline{z}(\alpha) \qquad \alpha = 1, 2, \cdots, k$$

choosing L_1 such that

$$L\Sigma L_1^{*\prime} = \emptyset \; .$$

Letting R^2 denote the exponent of Q, then

$$R^2 = \sum_{\alpha=1}^{k} n_\alpha\left(\bar{\underline{\xi}}(\alpha) - \underline{\mu}(\alpha)\right)^{*\prime}\Sigma^{-1}\left(\bar{\underline{\xi}}(\alpha) - \underline{\mu}(\alpha)\right)$$

$$= \sum_{\alpha=1}^{k} n_\alpha\left(L\bar{\underline{\xi}}(\alpha) - L\underline{\mu}(\alpha)\right)^{*\prime}(L\Sigma L^{*\prime})^{-1}\left(L\bar{\underline{\xi}}(\alpha) - L\underline{\mu}(\alpha)\right)$$

$$+ \sum_{\alpha=1}^{k} n_\alpha\left(L_1\bar{\underline{\xi}}(\alpha) - L_1\underline{\mu}(\alpha)\right)^{*\prime}(L_1\Sigma L_1^{*\prime})^{-1}\left(L_1\bar{\underline{\xi}}(\alpha) - L_1\underline{\mu}(\alpha)\right)$$

and under the assumption that the hypothesis is true, the above becomes

$$R^2 = \sum_{\alpha=1}^{k} n_\alpha\left(\underline{y}(\alpha) - \underline{\nu}\right)^{*\prime}(L\Sigma L^{*\prime})^{-1}\left(\underline{y}(\alpha) - \underline{\nu}\right)$$

$$+ \sum_{\alpha=1}^{k} n_\alpha\left(\underline{z}(\alpha) - L_1\underline{\mu}(\alpha)\right)^{*\prime}(L_1\Sigma L_1^{*\prime})^{-1}\left(\underline{z}(\alpha) - L_1\underline{\mu}(\alpha)\right)$$

and to maximize the likelihood, the exponent R^2 is minimized (observing that R^2 is real). Hence

$$\text{Min } R^2 = \text{Min } \sum_{=1}^{k} n_\alpha\left(\underline{y}(\alpha) - \underline{\nu}\right)^{*\prime}(L\Sigma L^{*\prime})^{-1}\left(\underline{y}(\alpha) - \underline{\nu}\right)$$

since differentiation with respect to $L_1\underline{\mu}(\alpha)$ gives $L_1\hat{\underline{\mu}}(\alpha) = L_1\bar{\underline{\xi}}(\alpha)$, so that the second term in R^2 vanishes. Thus

$$\text{Min } R^2 = \text{Min} \sum_{\alpha=1}^{k} n_\alpha [L\bar{\underline{\xi}}(\alpha) - \underline{\nu}]^{*\prime} (L\Sigma L^{*\prime})^{-1} [L\bar{\underline{\xi}}(\alpha) - \underline{\nu}]$$

$$= \text{Min} \sum_{\alpha=1}^{k} n_\alpha \operatorname{tr}(L\Sigma L^{*\prime})^{-1} [L\bar{\underline{\xi}}(\alpha) - \underline{\nu}][L\bar{\underline{\xi}}(\alpha) - \underline{\nu}]^{*\prime}$$

$$= \text{Min} \sum_{\alpha=1}^{k} n_\alpha \operatorname{tr}(L\Sigma L^{*\prime})^{-1} [L\bar{\underline{\xi}}(\alpha) - L\bar{\underline{\xi}} + L\bar{\underline{\xi}} - \underline{\nu}][L\bar{\underline{\xi}}(\alpha) - L\bar{\underline{\xi}} + L\bar{\underline{\xi}} - \underline{\nu}]^{*\prime}$$

where

$$\bar{\underline{\xi}} = \sum_{\alpha=1}^{k} \frac{n_\alpha \bar{\underline{\xi}}(\alpha)}{\sum_{\alpha=1}^{k} n_\alpha} .$$

Thus

$$\text{Min } R^2 = \text{Min } \operatorname{tr}(L\Sigma L^{*\prime})^{-1} \cdot \left\{ \sum_{\alpha=1}^{k} n_\alpha (L\bar{\underline{\xi}} - \underline{\nu})(L\bar{\underline{\xi}} - \underline{\nu})^{*\prime} + \sum_{\alpha=1}^{k} n_\alpha [L\bar{\underline{\xi}}(\alpha) - L\bar{\underline{\xi}}][L\bar{\underline{\xi}}(\alpha) - L\bar{\underline{\xi}}]^{*\prime} \right\}$$

and minimizing this with respect to $\underline{\nu}$, i.e., $\hat{\underline{\nu}} = L\bar{\underline{\xi}}$, then

$$\text{Min } R^2 = \underset{(\text{w.r.t.} L)}{\text{Min}} \operatorname{tr}(L\Sigma L^{*\prime})^{-1} L \left\{ \sum_{\alpha=1}^{k} n_\alpha [\bar{\underline{\xi}}(\alpha) - \bar{\underline{\xi}}][\bar{\underline{\xi}}(\alpha) - \bar{\underline{\xi}}]^{*\prime} \right\} L^{*\prime}$$

$$= \underset{(\text{w.r.t.} L)}{\text{Min}} \operatorname{tr}(L\Sigma L^{*\prime})^{-1} (LBL^{*\prime})$$

where B is the between groups s.s. and s.p. matrix from the CMANOVA table.

Now let [illegible] $^{1/2}$ then $TT^{*\prime}$ [illegible] and

$$\text{Min } R^2 = \text{Min tr}(TT^{*\prime})(\Sigma^{-1/2}B\Sigma^{-1/2}T^{*\prime})$$

$$= \text{Min tr } Y(\Sigma^{-1/2}B\Sigma^{-1/2})Y^{*\prime}$$

where $Y = (TT^{*\prime})^{-1/2}T$, observing $YY^{*\prime} = I$. So the problem of minimizing R^2 with respect to Λ reduces to minimizing the $\text{tr}[Y(\Sigma^{-1/2}B\Sigma^{-1/2})Y^{*\prime}]$ subject to the condition that $YY^{*\prime} = I$. As in the case of real positive definite symmetric matrices, the minimum is given by the sum of the smallest s roots of $\Sigma^{-1/2}B\Sigma^{-1/2}$ [Rao(1965)], i.e., the s smallest roots of

$$|\Sigma^{-1/2}B\Sigma^{-1/2} - \lambda_i I| = 0 \qquad i = 1, 2, \cdots, s$$

or

$$|B - \lambda_i \Sigma| = 0 \qquad i = 1, 2, \cdots, s \quad .$$

Thus the maximum value of Q under H is given as

(constant) $\cdot$ exp(− sum of the smallest s roots of $\Sigma^{-1/2}B\Sigma^{-1/2}$)

and the likelihood ratio statistic λ is given as

$$\lambda = e^{-(\text{sum of the smallest s roots of } \Sigma^{-1/2}B\Sigma^{-1/2})}$$

and

$$-2 \ln \lambda = (\text{sum of s smallest roots of } \Sigma^{-1/2}B\Sigma^{-1/2})$$

$$= 2(\lambda_{r+1} + \lambda_{r+2} + \cdots + \lambda_p)$$

[illegible] $^{-1/2}B^{-1/2}$.
The distribution of $-2 \ln \lambda$ is an asymptotic chi-square with degrees of freedom given by the number of parameters specified in Ω minus the number of restrictions of the parameters specified by the hypothesis.

Rao (1965), in the real case, has presented a geometrical argument for determining the degrees of freedom for this particular test. Wani and Kabe (1970) and Kshirsagar (1971) have presented an analytical argument which is somewhat easier to follow for the non-geometrician. Kshirsagar's procedure will be followed here to determine the degrees of freedom for the chi-square test of dimensionality.

The number of parameters specified in Ω is given by 2pk. The number of parameters restricted by the maximization of Q under H needs to be determined. Recall the procedure for maximizing Q under H. First the transformation by L and L_1 took

$$\underline{\mu}(\alpha) \rightarrow \begin{cases} L\underline{\mu}(\alpha) & \alpha = 1, 2, \cdots, k \\ L_1\underline{\mu}(\alpha) & \end{cases}$$

and the exponent R^2 was minimized, where R^2 was given as

$$R^2 = \sum_{\alpha=1}^{k} n_\alpha [\underline{y}(\alpha) - \underline{\nu}]^{*\prime}(L\Sigma L^{*\prime})^{-1}[\underline{y}(\alpha) - \underline{\nu}]$$

$$+ \sum_{\alpha=1}^{k} n_\alpha [\underline{z}(\alpha) - L_1\underline{\mu}(\alpha)]^{*\prime}(L_1\Sigma L_1^{*\prime})^{-1}[\underline{z}(\alpha) - L_1\underline{\mu}(\alpha)]$$

= Part I + Part II .

This minimization was done in steps:

Step I: Min of Part II, i.e., min of Part II with respect to $L_1\underset{\sim}{\mu}(\alpha)$. The number of parameters restricted was $2(p-s)k$ or $2rs$.

Step II: Min of Part I, i.e., min of Part I with respect to $\underset{\sim}{\nu}$. The number of parameters restricted was $2s$.

This left

$$\underset{(\text{w.r.t.} L)}{\text{Min}\ R^2} = \text{Min tr}(L\Sigma L^{*\prime})^{-1}(LBL^{*\prime}) \quad .$$

Now L is an $s \times p$ matrix of rank s and the number of parameters restricted needs to be determined. Write

$$L = s[\underset{s}{H} \mid \underset{p-s}{G}]$$

and note that

$$H^{-1}L = [I_s \mid H^{-1}G] \quad .$$

Returning to

$$\text{Min tr}(L\Sigma L^{*\prime})^{-1}(LBL^{*\prime})$$

which can be written as

$$\text{Min tr}[L^{*\prime}(L\Sigma L^{*\prime})^{-1}(LB)]$$

or as

$$[illegible]\ _{(p)};$$

where [illegible] is any [illegible] transformation matrix. Choose P to be H^{-1}, thus

$$[illegible] = [I_s \mid L_0] \quad \text{where} \quad L_0 = P^{-1}G$$

hence the minimum involves only functions of L_0, where L_0 has dimensionality given by $s \times (p-s)$. Thus only $2[s \times (p-s)]$ parameters are restricted in the minimization of [illegible]. Therefore the number of parameters restricted under H is given as

$$2rk + 2s + 2s(p-s)$$

and the degrees of freedom of the chi-square test is given by

$$d.f. = 2pk - [2rk + 2s + 2s(p-s)]$$

$$= 2(p-r)(q-r)$$

where $q = k+1$.

In summary, the test of the hypothesis that the number of linear discriminant functions is r, or the number of covariates is s = p-r, is given by the likelihood ratio criterion as

$$-2 \ln [illegible]_{r+1} + [illegible]_{r+2} + \cdots + \lambda_p) \sim \chi^2_{2(p-r)(q-r)}$$

where $\lambda_{r+1}, \lambda_{r+2}, \cdots, \lambda_p$ are the s smallest roots of

$$|B - \lambda [illegible]| = 0$$

where B is the between groups s.s. and s.p. matrix obtained from the CMANOVA table.

If Σ is unknown, then the estimate, $A/n-q$, where A is the within s.s. and s.p. from the CMANOVA table, may be used to obtain an approximate test for the number of discriminant functions needed. The test will be given as the s smallest roots of

$$\left| B - \frac{\hat{\lambda}}{n-q} A \right| = |B - \eta A| = 0 \quad .$$

The test statistic is given by

$$(n-q)(2)(\eta_{r+1} + \eta_{r+2} + \cdots + \eta_p) \mathrel{\dot{\sim}} \chi^2_{2(p-r)(q-r)}$$

where $\eta_{r+1}, \eta_{r+2}, \ldots, \eta_p$ are the s smallest roots of

$$|B - \eta A| = 0 \quad .$$

Instead of considering the roots of $|B - \eta A| = 0$, consider the roots of

$$|B - r^2(A+B)| = 0$$

or

$$\left| B - \frac{r^2}{1-r^2} A \right| = 0 \quad .$$

i.e.,

$$\eta = \frac{r^2}{1-r^2} \quad \text{or} \quad \frac{\hat{\lambda}}{n-q} = \frac{r^2}{1-r^2}$$

or

$$1 + \frac{\hat{\lambda}}{n} = \frac{1}{1-r^2}$$

for large n. Using the substitution

$$\log\left(1 + \frac{\hat{\lambda}}{n}\right) = -\log(1-r^2)$$

then

$$\frac{\hat{\lambda}}{n} \doteq -\log(1-r^2)$$

and

$$\frac{\hat{\lambda}_{r+1} + \hat{\lambda}_{r+2} + \cdots + \hat{\lambda}_p}{n} = -\log \prod_{i=r+1}^{p} (1-r_i^2)$$

or

$$-2n \log \prod_{i=r+1}^{p} (1-r_i^2) \mathrel{\dot{\sim}} \chi^2_{2(p-r)(q-r)}$$

where the r_i^2 are the roots of $|B - r(A+B)| = 0$. This gives a procedure for estimating the dimensionality of the mean space. The procedure is to test sequentially the following hypotheses.

	dimensionality	space	criterion
H: r=0	0	1 pt.	$-2n \log \prod_{i=1}^{p} (1-r_i^2) \mathrel{\dot{\sim}} \chi^2_{2pq}$
H: r=1	1	2 pts. (collinear)	$-2n \log \prod_{i=2}^{p} (1-r_i^2) \mathrel{\dot{\sim}} \chi^2_{2(p-1)(q-1)}$
H: r=2	2	3 pts. (coplanar)	$-2n \log \prod_{i=3}^{p} (1-r_i^2) \mathrel{\dot{\sim}} \chi^2_{2(p-2)(q-2)}$
.			
.			
.			

If the means are known, then the number of discriminant functions can be determined exactly by examining the rank of $\bar{\mu}$ where $\bar{\mu}$ is defined as

$$\bar{\mu} = \mu\left(I - \frac{1}{k} E_{kk}\right) .$$

To see this, consider L as defined by (3.2.2) and under H, the hypothesis that the number of discriminant functions is r, then

$$L\mu = \underline{\nu}E_{1k} \quad .$$

Observe

$$L\bar{\mu} = L\mu\left(I - \frac{1}{k}E_{kk}\right) = \underline{\nu}E_{1k}\left(I - \frac{1}{k}E_{kk}\right) = \emptyset$$

or

$$L_{s\times p}\bar{\mu}_{p\times k} = \emptyset \quad . \tag{3.2.3}$$

Now the rank of L is s since $\underline{\ell}(1)$, $\underline{\ell}(2)$, $\cdots$, $\underline{\ell}(s)$ are linearly independent and this is the exact number of covariates that exist. From (3.2.3), we have

$$L_{s\times p}E_{p\times p} = 0 \tag{3.2.4}$$

where $E_{p\times p} = \bar{\mu}_{p\times k}\bar{\mu}^{*\prime}_{k\times p}$, and from (3.2.3) we have

$$\underline{\ell}^{*\prime}(1)E = 0$$

$$\underline{\ell}^{*\prime}(2)E = 0$$

$$\vdots$$

$$\underline{\ell}^{*\prime}(s)E = 0$$

and there are no more non-trivial solutions to the system $E\underline{t} = 0$, thus we have the $r(E) = p-s$ which is the same as $r(\bar{\mu}) = p-s = r$.

Thus the rank of $\bar{\mu}$ is the same as the number of discriminant functions needed. Attention should be called to the fact that this is the exact number of discriminate functions needed and not an estimate.

In the first section, Wilks' Λ was defined as

$$ {}_c\Lambda = \frac{|A|}{|A+B|} $$

and it was shown that the test statistic for the hypothesis of equal means, i.e., $\underline{\mu}(1) = \underline{\mu}(2) = \cdots = \underline{\mu}(k)$, was given by

$$ -m \log {}_c\Lambda \overset{\cdot}{\sim} \chi^2_{2pq} \quad . \tag{3.2.5} $$

The distribution and correction factor $m = 2n - p - q$ were derived by Khatri (1965).

When the dimensionality of the mean space is zero, this is the same as saying all of the means are equal and the test was given as

$$ -2n \log \prod_{i=1}^{p} (1-r_i^2) \overset{\cdot}{\sim} \chi^2_{2pq} \tag{3.2.6} $$

where the r_i^2 are obtained from $|B - r^2(A+B)| = 0$. But the hypothesis that the dimensionality is zero is the same as the hypothesis that $\underline{\mu}(1) = \underline{\mu}(2) = \cdots = \underline{\mu}(k)$. Thus the two test statistic (3.2.5) and (3.2.6) should be equivalent. Since

$$ {}_c\Lambda = \frac{|A|}{|A+B|} = \prod_{i=1}^{p} (1-r_i^2) $$

it is seen that the two tests differ only in the constants $m = 2n - p - q$ and $2n$. Khatri's result has been corrected, so it will be used in (3.2.6) and the correction will be adjusted in general without a formal proof. Thus a test for dimensionality being r will be given as

$$[2n - (p-r) - (q-r)] \log \prod_{i=r+1}^{p} (1-r_i^2) \sim \chi^2_{2(p-r)(q-r)} .$$

As stated before, the hypothesis that the dimensionality is r is equivalent to the hypothesis of $L_{s\times p}\underline{\mu}(\alpha) = \underline{\nu}$, for $\alpha = 1, 2, \cdots, k$, which is also equivalent to the hypothesis that $r(L) = s = p-r$. Now the matrix L and the vector $\underline{\nu}$ are unspecified, and for a new observation $\underline{\xi}$ from π_α $(\alpha = 1, 2, \cdots, k)$, $L\underline{\xi}$ would provide the s covariates, $\underline{\ell}^{*\prime}(1)\underline{\xi}$, $\underline{\ell}^{*\prime}(2)\underline{\xi}, \cdots, \underline{\ell}^{*\prime}(s)\underline{\xi}$. Now these s variables all have the same mean in π_α $(\alpha = 1, 2, \cdots, k)$ and are no good as discriminators. Consider now $L_{r\times p}\underline{\xi}$ such that $L\Sigma L_1^{*\prime} = \emptyset$. Now these r variables are uncorrelated with $L\underline{\xi}$ and these can be used as discriminators. Unfortunately neither L nor L_1 is known, but the maximum likelihood estimator of L can be obtained.

In deriving the likelihood ratio procedure, the

$$\text{Min } \operatorname{tr}(L\Sigma L^{*\prime})^{-1}(LBL^{*\prime})$$

was considered. Now the maximum likelihood estimate of L would be the matrix L that does minimize the above. Recall that $T = L\Sigma^{-1/2}$ and the above became

$$\underset{T}{\text{Min}} \operatorname{tr}(TT^{*\prime})^{-1}(T\Sigma^{1/2}B\Sigma^{1/2}T^{*\prime})$$

and then set $Y = (TT^{*\prime})^{1/2}T$, and obtained

$$\underset{Y_{s\times p}}{\text{Min}} \operatorname{tr}[Y(\Sigma^{-1/2}B\Sigma^{-1/2})Y^{*\prime}]$$

subject to $YY^{*\prime} = I$.

To find the solution, consider the last s eigenvectors $\underline{Y}$ satisfying

$$(\Sigma^{-1/2}B\Sigma^{-1/2} - \lambda\ I)\underline{Y}_{p\times 1} = 0 \tag{3.2.7}$$

corresponding to the s smallest eigenvalues of

$$|\Sigma^{-1/2}B\Sigma^{-1/2} - \lambda\ I| = 0\ . \tag{3.2.8}$$

The condition $YY^{*\prime} = I$ will be satisfied since the eigenvalues corresponding to distinct roots are orthogonal and eigenvalues corresponding to repeated roots can be made orthogonal. We can take $\hat{T}^{*\prime}$ to be $\hat{Y}^{*\prime}$ from the s orthogonal eigenvectors $[\underline{Y}_1, \underline{Y}_2, \cdots , \underline{Y}_s]$ corresponding to the s smallest roots and

$$\hat{L}^{*\prime} = \Sigma^{-1/2}\hat{T}^{*\prime} = \Sigma^{-1/2}\hat{Y}^{*\prime}$$

and this is the maximum likelihood estimator of $L^{*\prime}$.

If we consider the largest p-s eigenvectors corresponding (3.2.7) and (3.2.8), these will provide the r discriminant functions. To show that $L\Sigma L_1^{*\prime} = \emptyset$, consider

$$(\Sigma^{-1/2}B\Sigma^{-1/2} - \lambda\ I)\underline{Y} = 0$$

or

$$(B\Sigma^{-1/2} - \lambda\ \Sigma^{1/2}\ I)\underline{Y} = 0$$

or

$$(B - \lambda\ \Sigma\ I)\Sigma^{-1/2}\underline{Y} = 0$$

or

$$(B - \lambda\ \Sigma\ I)\underline{\ell} = 0$$

where $\underline{\ell} = \Sigma^{-1/2}\underline{Y}$. Now the $\underline{Y}$'s were mutually orthogonal, i.e.,

$$\underset{\sim}{Y}^{*\prime}(i)\underset{\sim}{Y}(j) = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

but $\underset{\sim}{\ell} = \Sigma^{-1/2}\underset{\sim}{Y}$, thus $\underset{\sim}{\ell}^{*\prime}(i)\Sigma\underset{\sim}{\ell}(j) = 0$, which implies that $L\Sigma L_1' = \emptyset$.

Now consider a new observation ξ from one of the populations π_α $(\alpha = 1, 2, \cdots, k)$. It is desired to place this observation into one of π_α $(\alpha = 1, 2, \cdots, k)$. The procedure for this discrimination process is as follows:

Step 1: Find the matrices B and A by finding the CMANOVA from the sample observations.

Step 2: Obtain roots of

$$|B - r^2(A+B)| = 0 \qquad r_1^2 > r_2^2 > \cdots > r_p^2 \quad .$$

Step 3: Determine the dimensionality sequentially by finding the smallest value of r such that

$$-m \log \prod_{i=r+1}^{p} (1-r_i^2) \mathrel{\dot\sim} \chi^2_{2(p-r)(q-r)}$$

is insignificant.

Step 4: Determine the eigenvectors

$$[B - r_i^2(A+B)]\underset{\sim}{\ell}(i) = 0 \quad i = 1, 2, \cdots, r$$

corresponding to $r_1^2, \cdots, r_r^2$.

Step 5: Normalize the $\underset{\sim}{\ell}(i)$ by

$$\underset{\sim}{\ell}^{*\prime}(i)\Sigma\underset{\sim}{\ell}(j) = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}$$

if Σ is not known use $\hat{\Sigma} = A/n-q$.

Step 6: Then the discriminant functions are

$$\underline{\ell}^{*\prime}(1)\underline{\xi} = Y_1$$
$$\underline{\ell}^{*\prime}(2)\underline{\xi} = Y_2$$
$$\vdots$$
$$\underline{\ell}^{*\prime}(r)\underline{\xi} = Y_r \quad .$$

If a new observation is observed and the measured characteristics are recorded as $\underline{\xi}_0$, then

$$L_1\underline{\xi}_0 \quad \begin{bmatrix} \underline{\ell}^{*\prime}(1)\underline{\xi}_0 \\ \vdots \\ \underline{\ell}^{*\prime}(r)\underline{\xi}_0 \end{bmatrix} \quad .$$

The distance (modulus) of this point from the estimated mean of π_α is the quantity that will determine which population that $\underline{\xi}_0$ will go into, i.e.,

$$d^2_{\alpha,\underline{\xi}_0} = [\underline{\ell}^{*\prime}(1)\underline{\xi}_0-\underline{\ell}^{*\prime}(1)\bar{\underline{\xi}}(\alpha)]^{*\prime}[\underline{\ell}^{*\prime}(1)\underline{\xi}_0-\underline{\ell}^{*\prime}(1)\bar{\underline{\xi}}(\alpha)]$$
$$+ \cdots + [\underline{\ell}^{*\prime}(r)\underline{\xi}_0-\underline{\ell}^{*\prime}(r)\bar{\underline{\xi}}(\alpha)]^{*\prime}[\underline{\ell}^{*\prime}(r)\underline{\xi}_0-\underline{\ell}^{*\prime}(r)\bar{\underline{\xi}}(\alpha)]$$
$$= [L_1\underline{\xi}_0 - L_1\bar{\underline{\xi}}(\alpha)]^{*\prime}[L_1\underline{\xi}_0 - L_1\bar{\underline{\xi}}(\alpha)] \quad .$$

Step 7: Determine the minimum of

$$\left\{d^2_{1,\underline{\xi}_0}, d^2_{2,\underline{\xi}_0}, \cdots, d^2_{k,\underline{\xi}_0}\right\}$$

and assign π_j if minimum is $d^2_{j,\underline{\xi}_0}$. In case of a tie, one could randomize.

3. Goodness of Fit of a Single Hypothetical Discriminant Function

In the last section, a test of the hypothesis that r discriminant functions are needed to discriminate among the $k = q+1$ complex p-variate normal populations was developed. The following discussion will center around the case were the dimensionality is one, i.e., only one discriminant function is needed and with tests about this single discriminant function. In this section a test of the goodness of fit of a single hypothetical function will be developed. This hypothesis will consist of two parts, (i) whether a single hypothetical function is adequate, and (ii) whether the hypothetical function agrees with the true discriminant function. As in the real case, the hypothesis will be stated as the collinearity aspect and the direction aspect. The test statistic will be partitioned into two parts, one corresponding to the collinearity aspect and the other to the direction aspect. Independence between these factors will be shown.

The CMANOVA, as presented in the last sections, can be considered to be a regression of the complex vector $\underline{\xi}$ on the real dummy vector variable $\underline{Y}$. The dummy vector variables, $\underline{Y}_1, \underline{Y}_2, \cdots, \underline{Y}_q$, have components of either 0 or 1 depending on which population the $\underline{\xi}$ comes from, i.e.,

$$\underline{Y}'_\alpha = [\underbrace{0\ 0 \cdots 0}_{n_1} \cdots \underbrace{1\ 1 \cdots 1}_{n_\alpha} \cdots \underbrace{0\ 0 \cdots 0}_{n_k}]$$

and the regression can be written as

$$E(\underline{\xi}) = \underline{\mu}_\alpha \quad \text{if} \quad \underline{\xi} \in \pi_\alpha \quad \alpha = 1, 2, \cdots, k$$

$$= \underline{\mu}_k + (\underline{\mu}_1 - \underline{\mu}_k)\underline{Y}_1 + \cdots + (\underline{\mu}_q - \underline{\mu}_k)\underline{Y}_q$$

which can be rewritten as

$$E(X) = \underline{\mu}_k E_{1N} + \beta Y$$

where

$$X = [\underline{\xi}_{11}\ \underline{\xi}_{12} \cdots \underline{\xi}_{1n_1} | \cdots | \underline{\xi}_{k1}\ \underline{\xi}_{k2} \cdots \underline{\xi}_{kn_k}]$$

and

$$Y_{N\times q} = [\underline{Y}_1, \underline{Y}_2, \cdots , \underline{Y}_q]$$

with

$$\beta = [\underline{\mu}_1 - \underline{\mu}_k, \underline{\mu}_2 - \underline{\mu}_k , \cdots , \underline{\mu}_q - \underline{\mu}_k] .$$

With this in mind, the CMANOVA, may be written as

Source of Variation	d.f.	s.s. and s.p.
Regression (Between)	$\underline{q}$	$B = C_{xy}C_{yy}^{-1}C_{yx}$
Error (Within)	n-q	$A = C_{xx\ y} = C_{xx} - C_{xy}C_{yy}^{-1}C_{yx}$
TOTAL	n	$A+B = C_{xx}$

where for $J = (I_N - \frac{1}{N} E_{NN})$, then

$$C_{yy} = Y'JY \qquad C_{yx} = Y'JX^{*\prime}$$
$$C_{xy} = XJY \qquad C_{xx} = XJX^{*\prime} .$$

Define A_i as the matrix consisting of the first i rows and columns of A and likewise for B. With $A_0 = 1$ and $B_0 = 1$. Then

$$c^{\Lambda} = \frac{|A|}{|A+B|} = \prod_{i=1}^{p} \frac{\dfrac{|A_i|}{|A_{i-1}|}}{\dfrac{|A_i+B_i|}{|A_{i-1}+B_{i-1}|}}$$

which can be written in terms of regression as

$$c^{\Lambda} = \prod_{i=1}^{p} \frac{\dfrac{|A_i|}{|A_{i-1}|}}{\dfrac{|A_i+B_i|}{|A_{i-1}+B_{i-1}|}} = \prod_{i=1}^{p} \frac{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1}, \underline{Y}')}}{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1})}}$$

where $R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1})}$ is the multiple coherence of ξ_i with $\xi_1, \xi_2, \cdots, \xi_{i-1}$. This fact follows from

$$\frac{\dfrac{|A_i|}{|A_{i-1}|}}{\dfrac{|A_i+B_i|}{|A_{i-1}+B_{i-1}|}} = \frac{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1}, \underline{Y}')}}{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1})}}$$

which is evident from (2.4.4). It can be shown that

$$\frac{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1}, \underline{Y}')}}{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1})}} \sim \frac{\chi^2_{2[n-(q+1-i)]}}{\chi^2_{2[n-(q+1-i)]} + \chi^2_{2q}} = B(n-q+1-i,\ q) .$$

Thus,

$$c^{\Lambda} = \prod_{i=1}^{p} \frac{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1}, \underline{Y}')}}{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1})}} = \prod_{i=1}^{p} u_i \qquad (3.3.1)$$

where $u_i \sim B(n-q+1-i, q)$. This is the result obtained by Khatri (1965), who has shown that the u_i $(i = 1, 2, \cdots, p)$ form an independent set of beta variables. Another definition of ${}_c\Lambda$ is given by equation (3.3.1), i.e., if a statistic ${}_c\Lambda$ is distributed as $\prod_{i=1}^{p} u_i$ where u_i $(i = 1, 2, \cdots, p)$ are independently distributed as $B(n-q+1-i, q)$, then ${}_c\Lambda$ is said to have a complex Wilks' Λ distribution with parameters n, p, and q, denoted by ${}_c\Lambda(n, p, q)$. Thus

$$
\begin{aligned}
{}_c\Lambda &= \prod_{i=1}^{p} u_i \\
&= u_1 \prod_{i=2}^{p} u_i \\
&= u_1 \; {}_c\Lambda(n-1, p-1, q) \\
&= u_1 \, u_2 \; {}_c\Lambda(n-2, p-2, q) \; .
\end{aligned}
$$

The hypothesis of equal means can be restated in terms of the regression of $\underline{\xi}$ on $\underline{Y}$ as one of no association between $\underline{\xi}$ and $\underline{Y}$. This hypothesis can then be broken into the following sub-hypotheses

H_1: ξ_1 has no association with $\underline{Y}$

H_2: ξ_2 (eliminating ξ_1) has no association with $\underline{Y}$

.

.

.

H_p: ξ_p (eliminating $\xi_1, \xi_2, \cdots, \xi_{p-1}$) has no association with $\underline{Y}$.

If the overall hypothesis of no association is rejected, then at least one of the H_i $(i = 1, 2, \cdots, p)$ is rejected.

Observe that the criterion for the sub-hypothesis is given by

$$\frac{\dfrac{|A_i|}{|A_{i-1}|}}{\dfrac{|A_i+B_i|}{|A_{i-1}+B_{i-1}|}} = \frac{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1}, \underline{Y}')}}{1 - R^2_{\xi_i(\xi_1, \xi_2, \cdots, \xi_{i-1})}} = u_i \ .$$

It is desired to develop a goodness of fit test for a hypothetical discriminant function. Denote this function by $\underline{\ell}^{*\prime}\underline{\xi} = Z_1$ and consider a non-singular transformation of $\underline{\ell} \to \underline{Z}$ by $L\underline{\xi} = \underline{Z}$ where the first row of L is $\underline{\ell}^{*\prime}$. It is now desired to test that Z_1 is the reason that the dimensionality is not zero. Considering the population means of the transformed variables, Z_1 is the one that differs from $Z_2, Z_3, \cdots, Z_p$. If this is true, then after eliminating Z_1 the set $Z_2, \cdots, Z_p$ will have no association with $\underline{Y}$.

Thus the test statistic is given by

$${}_c\tilde{\Lambda} = \frac{{}_c\bar{\Lambda}}{\bar{u}_1} = \prod_{i=2}^{p} \bar{u}_i$$

where ${}_c\bar{\Lambda} = |LAL^{*\prime}|/|L(A+B)L^{*\prime}|$, i.e., the transformed CMANOVA. But observe

$${}_c\bar{\Lambda} = \frac{|LAL^{*\prime}|}{|L(A+B)L^{*\prime}|} = \frac{|A|}{|A+B|} = {}_c\Lambda$$

thus the test statistic is given by

$${}_c\tilde{\Lambda} = \frac{{}_c\Lambda}{\bar{u}_1} = \frac{{}_c\Lambda}{\underline{\ell}^{*\prime}(A)\underline{\ell}/\underline{\ell}^{*\prime}(A+B)\underline{\ell}} \qquad (3.3.2)$$

but from

$$ {}_c\Lambda = \prod_{i=1}^{p} (1-r_i^2) $$

so the test statistic becomes

$$ \frac{\prod_{i=1}^{p} (1-r_i^2)}{\bar{u}_1} = \left\{ \frac{(1-r_1^2)}{\underline{\ell}^{*\prime} A \underline{\ell} / \underline{\ell}^{*\prime} (A+B) \underline{\ell}} \right\} \left\{ \prod_{i=2}^{p} (1-r_i^2) \right\} \qquad (3.3.3) $$

Observe that the second factor is the test statistic for the dimensionality of one. The first factor tests that the discriminant function is given by $\underline{\ell}^{*\prime}\underline{\xi}$. Observe further that the two factors are not independent, since the r_i^2 $(i = 1, 2, \cdots, p)$ are not independent. From (3.3.3) an approximate χ^2 test can be constructed.

Now a factorization of ${}_c\tilde{\Lambda}$ into two independent factors that measure the direction aspect and the partial collinearity aspect can be achieved or ${}_c\tilde{\Lambda}$ can be the factor into two independent factors measuring the collinearity and 'partial' direction aspect. The use of the word 'partial' will be discussed later. In the real and general case this factorization was developed by Bartlett (1951), who gave a geometrical argument. Kshirsagar (1970) gave an analytical argument for the same factorization. The work presented here in the complex case is a combination of the two, with extensive use of Kshirsagar's work. The factorization of ${}_c\Lambda$ into the direction and 'partial' collinearity aspects is presented. The alternate factorization can be achieved in a similar manner.

If the hypothesis that $\underline{\ell}^{*\prime}\underline{\xi}$ is the only discriminant function is true, then $\underline{\ell}^{*\prime}\underline{\xi}$ is the only variable that has a different mean among the k populations. Any other linear combination of $\underline{\xi}$, uncorrelated with $\underline{\ell}^{*\prime}\underline{\xi}$, has the same mean in all the k groups. Let L be a $p \times p$ non-singular matrix with $\underline{\ell}^{*\prime}$ as its first row and let L be such that $L\Sigma L^{*\prime} = I$. (There is no loss of generality for this assumption.) Transform from $\underline{\xi}$ to $\underline{Z}$, where $\underline{Z}' = [Z_1, Z_2, \cdots, Z_p]$ by

$$\underline{Z} = L\,\underline{\xi} \quad .$$

Then Z_1 is the hypothetical discriminant function and $Z_2, \cdots, Z_p$ are all uncorrelated with Z_1, as $L\Sigma L^{*\prime} = I$. Under the hypothesis, $Z_2, \cdots, Z_p$ all have the same mean in the k groups. The test for the hypothesis has been developed and was shown to be on (3.3.1)

$$_c\tilde{\Lambda} = \frac{_c\Lambda}{\underline{\ell}^{*\prime}A\underline{\ell}/\underline{\ell}^{*\prime}(A+B)\underline{\ell}} = \frac{_c\Lambda}{_c\Lambda_1} \quad . \qquad (3.3.4)$$

Now assuming that the factorization of $_c\tilde{\Lambda}$ can be achieved and $_c\tilde{\Lambda}$ can be represented as

$$_c\tilde{\Lambda} = {_c\Lambda_2} \cdot {_c\Lambda_3}$$

where $_c\Lambda_2$ is the direction factor and $_c\Lambda_3$ is the 'partial' collinearity factor.

The alternative factorization is given by

$$_c\tilde{\Lambda} = {_c\Lambda_4} \cdot {_c\Lambda_5}$$

where $_c\Lambda_4$ is the collinearity factor and $_c\Lambda_5$ is the 'partial' direction factor.

To show that the above is indeed the case, consider the following argument. Let D be the between groups matrix for the transformed variables Z, then

$$D = LBL^{*'} = [d_{ij}]$$

and the within groups matrix is W, given by

$$W = LAL^{*'} = [w_{ij}] \quad .$$

As in the real case, it is obvious that W has the Wishart density given by $CW_p(W|n-q|I)$ since $L\Sigma L^{*'} = I$ and likewise $D = LBL^{*'}$ is distributed as a non-central complex Wishart, with the non-centrality parameter due only to the difference in means of Z_1 alone in the k groups and thus affects b_{11} only, i.e., b_{11} is a non-central χ^2 with 2q degrees of freedom. So the density of D may be written as

$$CW_p(D|q|I)\S(b_{11})$$

where $\S(b_{11})$ is the contribution due to the non-centrality. An explicit expression for $\S(b_{11})$ will not be necessary in this development.

Now $W \sim CW_p(W|n-q|I)$ where $W = [w_{ij}]$. From this we have

$$w_{11}, \frac{w_{i1}}{\sqrt{w_{11}}} \; (i = 2, \cdots, p) \quad \text{and} \quad \tilde{w}_{ij} = w_{ij} - \frac{w_{i1}w_{j1}}{w_{11}} \; (i,j = 2, \ldots, p)$$

are independently distributed. The $\frac{w_{i1}}{\sqrt{w_{11}}}$ are complex standard normal and independent variables, while the matrix $\tilde{W} = [\tilde{w}_{ij}]$ is a complex Wishart of order p-1 . This follows, as in the real case, from the fact that if $S \sim CW_p(S|n|I)$ and

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

then $S_{22\cdot 1} = S_{22} - S_{21}S_{11}^{-1}S_{12}$ is of order p-k and distributed as a complex Wishart and is independent of S_{11} and $S_{21}S_{11}^{-1}$. So applying this result, we have

$$\tilde{W} \sim CW_{p-1}(\tilde{W}|n-q-1|I) \quad .$$

In a similar matter, D can be split up as

$$d_{11}, \frac{d_{i1}}{\sqrt{d_{11}}} \ (i = 2, \cdots, p) \quad \text{and} \quad \tilde{d}_{ij} = d_{ij} - \frac{d_{i1}d_{j1}}{d_{11}} \ (i,j = 2, \cdots, p)$$

and are independently distributed. The $d_{i1}/(d_{11})^{1/2}$ are $CN_1(0, 1)$ and are independent and $\tilde{D} = [\tilde{d}_{ij}]$ is distributed as $CW_{p-1}(\tilde{D}|q-1|I)$. All of these elements are independent of W and hence $\tilde{W}$.

Consider the variable

$$\theta_i = \sqrt{\frac{d_{11}w_{11}}{d_{11}+w_{11}}}\left(\frac{d_{i1}}{d_{11}} - \frac{w_{i1}}{w_{11}}\right) \quad \text{for } i = 2, \cdots, p$$

which is a $CN_1(0, 1)$ random variable. To see this consider the conditional density of θ_i for d_{11} and w_{11} fixed. This is $CN_1(0, 1)$ but does not depend on w_{11} or b_{11}, hence it is the unconditional distributions. Let $\underline{\theta}$ denote the column vectors θ_i $(i = 2, 3, \cdots, p)$.

Thus the following results are obtained

$$\tilde{W} \sim CW_{p-1}(\tilde{W}|n-q-1|I)$$
$$\tilde{D} \sim CW_{p-1}(\tilde{D}|q-1|I)$$
$$\theta \sim CN_{p-1}(\underline{0}, I)$$

and all are independent. Observe that $\underset{\sim}{\theta}\,\underset{\sim}{\theta}^{*\prime}$ is distributed as a pseudo-complex Wishart with one degree of freedom.

Recall that Wilks' ${}_c\Lambda$ with parameters n, p, q was defined as

$$ {}_c\Lambda = \frac{|A|}{|A+B|} = {}_c\Lambda(n,\ p,\ q) $$

where $A \sim CW_p(A|n-q|\Sigma)$ independent of $B \sim CW_p(B|q|\Sigma)$. The parameter n is the d.f. of A plus the degrees of freedom of B; p is the order of A+B; q is the degrees of freedom of B. Using this terminology, then

$$ \frac{|\tilde{W}|}{|\tilde{W}+\theta\ \theta^{*\prime}|} \sim {}_c\Lambda(n-q-1+1,\ p-1,\ 1) = {}_c\Lambda_5 $$

$$ \frac{|\tilde{W}+\theta\ \theta^{*\prime}|}{|\tilde{W}+\tilde{D}+\theta\ \theta^{*\prime}|} \sim {}_c\Lambda(n-1,\ p-1,\ q-1) = {}_c\Lambda_4 $$

$$ \frac{|\tilde{W}|}{|\tilde{W}+\tilde{D}|} \sim {}_c\Lambda(n-1,\ p-1,\ q-1) = {}_c\Lambda_3 $$

$$ \frac{|\tilde{W}+\tilde{D}|}{|\tilde{W}+\tilde{D}+\theta\ \theta^{*\prime}|} \sim {}_c\Lambda(n-1,\ p-1,\ 1) = {}_c\Lambda_2 \quad . $$

Furthermore, independence between ${}_c\Lambda_2$ and ${}_c\Lambda_3$ is obtained or independence between ${}_c\Lambda_4$ and ${}_c\Lambda_5$ is obtained. The same argument for the real case is valid in the complex situation, i.e., if $A \sim CW_p(A|f|I)$ and $\underset{\sim}{\xi}_i \sim CN_p(\underset{\sim}{0},\ I)$ $(i = 1,\ 2,\ \cdots,\ n)$, then

$$ t = \frac{|A|}{\left|A + \sum_{i=1}^{n} \underset{\sim}{\xi}_i \underset{\sim}{\xi}_i^{*\prime}\right|} $$

is independently distributed of $A + \sum_{i=1}^{n} \underline{\xi}_i \underline{\xi}_i^{*'}$.

Now, it must be shown that ${}_c\tilde{\Lambda}$ is indeed given by the product ${}_c\Lambda_2 \cdot {}_c\Lambda_3$. And then it will be shown that ${}_c\Lambda_2$ is the direction factor and ${}_c\Lambda_3$ is the 'partial' collinearity factor. The alternate factorization, ${}_c\Lambda = {}_c\Lambda_4 \cdot {}_c\Lambda_5$, can also be shown. Consider

$${}_c\Lambda_2 \cdot {}_c\Lambda_3 = \frac{|\tilde{W}|}{|\tilde{W}+\tilde{D}+\theta\ \theta^{*'}|}$$

$$= \frac{\frac{|W|}{w_{11}}}{|\tilde{W}+\tilde{D}+\theta\ \theta^{*'}|}$$

and observe that

$$|\tilde{W} + \tilde{D} + \theta\ \theta^{*'}| = \frac{1}{b_{11}+w_{11}} |W+D| \tag{3.3.5}$$

then the above can be written as

$$= \frac{\frac{|W|}{w_{11}}}{\frac{|W+D|}{b_{11}+w_{11}}}$$

$$= \frac{[\underline{\ell}^{*'}(A+B)\underline{\ell}]}{(\underline{\ell}^{*'}A\underline{\ell})} \cdot \frac{|LAL^{*'}|}{|L(A+B)L^{*'}|} .$$

Since $w_{11} = \underline{\ell}^{*'}A\underline{\ell}$ and $(w_{11} + b_{11}) = \underline{\ell}^{*'}(A+B)\underline{\ell}$

$$= \frac{\dfrac{|\tilde{A}|}{|A+B|}}{\dfrac{\underline{\ell}^{*\prime} A \underline{\ell}}{\underline{\ell}^{*\prime} (A+B) \underline{\ell}}}$$

$$= \frac{{}_c\Lambda}{{}_c\Lambda_1}$$

and it is seen that (3.3.4) is verified.

Now it must be shown that ${}_c\Lambda_2$ is the direction factor and ${}_c\Lambda_3$ is the 'partial' collinearity factor. Consider ${}_c\Lambda_2$ first. To see this, define a matrix H by

$$\left[\begin{array}{c|ccc} w_{11} & w_{12} & \cdots & w_{1p} \\ \hline w_{21} & & & \\ & & w_{ij} + \tilde{d}_{ij} & \\ & & (i,j = 2,\cdots,p) & \\ w_{p1} & & & \end{array}\right]$$

and a vector k by

$$\underline{k}' = [\sqrt{d_{11}}\ ,\ d_{12}/\sqrt{b_{11}}\ ,\ \cdots\ ,\ d_{1p}]$$

then observe that

$$D + W = H + \underline{\lambda}\ \underline{k}^{*\prime}$$

and

$$|H| = w_{11}|\tilde{W} + \tilde{D}| \qquad (3.3.6)$$

and from (3.3.5), that

$$|H + \underline{k}\ \underline{k}^{*\prime}| = (w_{11}+b_{11})\,|W + \tilde{D} + \underline{\theta}\ \underline{\theta}^{*\prime}| \quad . \tag{3.3.7}$$

Then ${}_c\Lambda_2$ may be written in terms of (3.3.6) and (3.3.7) as

$${}_c\Lambda_2 = \frac{(b_{11}+w_{11})}{w_{11}} \cdot \frac{|H|}{|H + \underline{k}\ \underline{k}^{*\prime}|}$$

then the above may be written as

$$= \frac{(b_{11}+w_{11})}{(w_{11})} \cdot \frac{1}{1 + \underline{k}^{*\prime}H^{-1}\underline{k}} \quad .$$

Since $|H + \underline{k}\ \underline{k}^{*\prime}| = |H| \cdot |I + H^{-1}\underline{k}\ \underline{k}^{*\prime}| = |H| \cdot (1 + \underline{k}^{*\prime}H^{-1}\underline{k})$ and ${}_c\Lambda_2$ can be further simplified to

$$= \frac{(b_{11}+w_{11})}{(w_{11})} \cdot [1 - \underline{k}^{*\prime}(H + \underline{k}\ \underline{k}^{*\prime})^{-1}\underline{k}]$$

this can be seen by using the fact that

$$(I + PQ)^{-1} = I - P(I + QP)^{-1}Q \quad .$$

Noting that

$$\underline{k}^{*\prime} = \underline{\ell}^{*\prime}L^{-1}DL^{*\prime -1}L^{*\prime}/[\underline{\ell}^{*\prime}(L^{-1}DL^{*\prime -1})\underline{\ell}]^{1/2}$$

then ${}_c\Lambda_2$ may be written as

$${}_c\Lambda_2 = \frac{(b_{11}+w_{11})}{(w_{11})} \left\{ 1 - \frac{\underline{\ell}^{*\prime}L^{-1}DL^{*\prime}[L^{-1}(W+D)L^{*\prime -1}]^{-1}L^{-1}DL^{*\prime -1}\underline{\ell}}{\underline{\ell}^{*\prime}(L^{-1}DL^{*\prime -1})\underline{\ell}} \right\}$$

and observing that

$$\frac{b_{11}+w_{11}}{w_{11}} = \frac{\underline{\ell}^{*\prime}[L^{-1}(W+D)L^{*\prime\,-1}]\underline{\ell}}{\underline{\ell}^{*\prime}(L^{-1}WL^{*\prime\,-1})\underline{\ell}}$$

then

$$_c\Lambda_2 = \frac{1 - \dfrac{\underline{\ell}^{*\prime}L^{-1}DL^{*\prime}[L^{-1}(W+D)L^{*\prime\,-1}]^{-1}L^{-1}DL^{*\prime\,-1}\underline{\ell}}{\underline{\ell}^{*\prime}(L^{-1}DL^{*\prime\,-1})\underline{\ell}}}{\underline{\ell}^{*\prime}(L^{-1}WL^{*\prime\,-1})\underline{\ell}/\underline{\ell}^{*\prime}[L^{-1}(D+W)L^{*\prime\,-1}]\underline{\ell}}$$

and using the fact that $D = LBL^{*\prime}$ and $W = LAL^{*\prime}$, then

$$_c\Lambda_2 = \frac{1 - \underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}/\underline{\ell}^{*\prime}B\underline{\ell}}{\underline{\ell}^{*\prime}A\underline{\ell}/\underline{\ell}^{*\prime}(A+B)\underline{\ell}} \quad .$$

Then $_c\Lambda_3$ can be represented as

$$_c\Lambda_3 = \frac{_c\tilde{\Lambda}}{_c\Lambda_2}$$

$$= \frac{\dfrac{|A|}{|A+B|}}{1 - \underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}/\underline{\ell}^{*\prime}B\underline{\ell}} \quad .$$

Returning to the CMANOVA for the transformed variable $Z = L\underline{\xi}$, it can be expressed as

Source of Variation	d.f.	s.s. and s.p.
Regression (Between)	q	$LBL^{*\prime} = D = C_{zy}C_{yy}^{-1}C_{yz}$
Error (Within)	n-q	$LAL^{*\prime} = W = C_{zz\cdot y}$
TOTAL	n	$L(A+B)L^{*\prime} = C_{zz}$

The ${}_c\Lambda$ can be expressed as

$$ {}_c\Lambda = \frac{|C_{zz\cdot y}|}{|C_{zz}|} = \frac{|C_{xx\cdot y}|}{|C_{xx}|} \tag{3.3.8} $$

then using the fact that

$$ \begin{vmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{vmatrix} = |C_{xx}|\cdot|C_{yy} - C_{yx}C_{xx}^{-1}C_{xy}| = |C_{yy}|\cdot|C_{xx} - C_{xy}C_{yy}^{-1}C_{yx}| \tag{3.3.9} $$

it is seen that another expression for ${}_c\Lambda$ is given by

$$ {}_c\Lambda = \frac{|C_{yy\cdot x}|}{|C_{yy}|} \quad . \tag{3.3.10} $$

Now (3.3.9) is a measure of the "lack of relationship" between $L\underline{\xi}$ and $\underline{Y}$ and (3.3.10) is a measure of the "lack of relationship" between $\underline{Y}$ and $L\underline{\xi}$. Multiplying and dividing (3.3.10) by $|C_{yy\cdot t}|$, the residual of $\underline{Y}$ with the effect of $Z_1 = t$ removed, ${}_c\Lambda$ can be written as

$$ {}_c\Lambda = \frac{|C_{yy\cdot t}|}{|C_{yy}|} \cdot \frac{|C_{yy\cdot x}|}{|C_{yy\cdot t}|} \quad . $$

Consider the first factor, using (3.3.9), this can be written as

$$ \frac{|C_{yy\cdot t}|}{|C_{yy}|} = \frac{|C_{tt\cdot y}|}{|C_{tt}|} = \frac{|C_{tt} - C_{ty}C_{yy}^{-1}C_{yt}|}{|C_{tt}|} = \frac{\underline{\ell}^{*\prime}C_{xx\cdot y}\underline{\ell}}{\underline{\ell}^{*\prime}C_{xx}\underline{\ell}} = \frac{\underline{\ell}^{*\prime}A\underline{\ell}}{\underline{\ell}^{*\prime}(A+B)\underline{\ell}} \quad . $$

Hence another expression for ${}_c\tilde{\Lambda}$ is given by

$$ {}_c\tilde{\Lambda} = \frac{c\Lambda}{\underline{\ell}^{*\prime}A\underline{\ell}/\underline{\ell}^{*\prime}(A+B)\underline{\ell}} = \frac{c\Lambda}{|C_{yy\cdot t}|/|C_{yy}|} = \frac{|C_{yy\cdot x}|}{|C_{yy\cdot t}|} \tag{3.3.11} $$

which is the test statistic when $\underline{\ell}^{*\prime}\underline{\xi} = t$ is eliminated, thus measuring the association between $\underline{Y}$ and $\underline{Z}$, eliminating t . Now ${}_c\Lambda$ must be factored into ${}_c\Lambda_2$ and ${}_c\Lambda_3$. This will be achieved by factoring the numerator and denominator of (3.3.11).

Let § denote the sample projection of t on the $\underline{Y}$ space. The regression of $\underline{\xi}$ on $\underline{Y}$ is $C_{xy}C_{yy}^{-1}C_{yx}$, thus the regression of $t = \underline{\ell}^{*\prime}\underline{\xi}$ on $\underline{Y}$ is $\underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}\underline{Y}$. The sample projection § is given by $\S = \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}\underline{Y}$ or by $\S = m^{*\prime}\underline{Y}$ where $m^{*\prime} = \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}$. Let $\underline{\gamma}$ denote the remaining q-1 linear functions of the $\underline{Y}$'s . Then $|C_{yy\cdot z}|$ can be factored into $|C_{\S\S\cdot z}|\cdot|C_{\gamma\gamma\cdot\S z}|$. This follows from (3.3.9). Likewise $|C_{yy\cdot t}|$ may be factored into $|C_{yy\cdot t}| = |C_{\S\S\cdot t}|\cdot|C_{\gamma\gamma\cdot\S t}|$ and ${}_c\tilde{\Lambda}$ may be expressed as

$$ {}_c\tilde{\Lambda} = \frac{|C_{\S\S\cdot x}|\cdot|C_{\gamma\gamma\cdot\S x}|}{|C_{\S\S\cdot t}|\cdot|C_{\gamma\gamma\cdot\S t}|} $$

or as

$$ {}_c\tilde{\Lambda} = \left\{\frac{|C_{\S\S\cdot x}|}{|C_{\S\S\cdot t}|}\right\}\cdot\left\{\frac{|C_{\gamma\gamma\cdot\S\ x}|}{|C_{\gamma\gamma\cdot\S\ t}|}\right\} . $$

Now it will be shown that

$$ {}_c\Lambda_2 = \frac{|C_{\S\S\cdot x}|}{|C_{\S\S\cdot t}|} . $$

Now

$$ \begin{aligned} C_{\S\S\cdot x} &= m^{*\prime}C_{yy\cdot x}m \\ &= \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}[C_{yy} - C_{yx}C_{xx}^{-1}C_{xy}]C_{yy}^{-1}C_{yx}\underline{\ell} \\ &= \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}C_{yx}\underline{\ell} - \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}C_{yx}C_{xx}^{-1}C_{xy}C_{yy}^{-1}C_{yx}\underline{\ell} \\ &= \underline{\ell}^{*\prime}B\underline{\ell} - \underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell} \end{aligned} $$

and

$$C_{\S\S\cdot t} = m^{*\prime}C_{yy\cdot t}m$$

$$= \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}[C_{yy} - C_{yx}C_{tt}^{-1}C_{ty}]C_{yy}^{-1}C_{yx}\underline{\ell}$$

$$= \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}[C_{yy} - C_{yx}\underline{\ell}(\underline{\ell}^{*\prime}C_{xx}\underline{\ell})^{-1}\underline{\ell}^{*\prime}C_{xy}]C_{yy}^{-1}C_{yx}\underline{\ell}$$

$$= \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}[C_{yy} - (C_{yx}\underline{\ell}\ \underline{\ell}^{*\prime}C_{xy}/\underline{\ell}^{*\prime}C_{xx}\underline{\ell})]C_{yy}^{-1}C_{yx}\underline{\ell}$$

$$= \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}C_{yx}\underline{\ell} - [(\underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}C_{yx}\underline{\ell})^2/\underline{\ell}^{*\prime}C_{xx}\underline{\ell}]$$

$$= \underline{\ell}^{*\prime}B\underline{\ell} - [(\underline{\ell}^{*\prime}B\underline{\ell})^2/\underline{\ell}^{*\prime}(A+B)\underline{\ell}] \quad .$$

Thus

$$\frac{|C_{\S\S\cdot x}|}{|C_{\S\S\cdot t}|} = \frac{\underline{\ell}^{*\prime}B\underline{\ell} - \underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}}{\underline{\ell}^{*\prime}B\underline{\ell} - \dfrac{(\underline{\ell}^{*\prime}B\underline{\ell})^2}{\underline{\ell}^{*\prime}(A+B)\underline{\ell}}}$$

$$= \frac{\underline{\ell}^{*\prime}B\underline{\ell}\left\{1 - \dfrac{\underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}}{\underline{\ell}^{*\prime}B\underline{\ell}}\right\}}{\dfrac{\underline{\ell}^{*\prime}B\underline{\ell}\ \underline{\ell}^{*\prime}A\underline{\ell}}{\underline{\ell}^{*\prime}(A+B)\underline{\ell}}}$$

$$= \frac{1 - \dfrac{\underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}}{\underline{\ell}^{*\prime}A\underline{\ell}}}{\dfrac{\underline{\ell}^{*\prime}A\underline{\ell}}{\underline{\ell}^{*\prime}(A+B)\underline{\ell}}} = {}_c\Lambda_2 \quad .$$

In a similar matter, ${}_c\Lambda_3$ follows. Thus we have

$${}_c\Lambda_2 = \frac{|C_{\S\S\cdot x}|}{|C_{\S\S\cdot t}|} \quad \text{and} \quad {}_c\Lambda_3 = \frac{|C_{\gamma\gamma\cdot\S x}|}{|C_{\gamma\gamma\cdot\S t}|} \quad .$$

Now ${}_c\Lambda_2$ is a measure of the association of $\S$, the sample projection on the $\underline{Y}$ space, with the $\underline{\xi}$ space with the effect of the discriminant function, t, removed. Now $\S = m^{*\prime}\underline{Y}$, where $\underline{m}^{*\prime} = \underline{\ell}^{*\prime}C_{xy}C_{yy}^{-1}$. Observe that $\underline{m}$ is determined by $\underline{\ell}$, the given direction, thus one would expect no association to be present if $\underline{\ell}^{*\prime}\underline{\xi} = t$ is eliminated. Since $\underline{Y} = \underline{m}^{*\prime}\underline{Y}$ is considered first and then the $\underline{\gamma}$'s are considered after eliminating $\underline{m}^{*\prime}\underline{Y}$, ${}_c\Lambda_3$ is called the 'partial' collinearity factor. Considering the γ's first and then considering $\S$ after the elimination of γ will give the alternate factorization, i.e.,

$$ {}_c\tilde{\Lambda} = {}_c\Lambda_4 \cdot {}_c\Lambda_5 $$

where ${}_c\Lambda_4$ is the collinearity factor and ${}_c\Lambda_5$ is the 'partial' direction factor.

${}_c\Lambda_3$ is the collinearity factor since it depends on the number, but not the direction, of the remaining q-1 variables.

In summary, we have the test for the hypothesis that $\underline{\ell}^{*\prime}\underline{\xi}$ is the only linear discriminant function needed to discriminate among the k = q+1 complex p-variate normal population based on the statistic

$$ {}_c\tilde{\Lambda} = \frac{{}_c\Lambda}{\dfrac{\underline{\ell}^{*\prime}A\underline{\ell}}{\underline{\ell}^{*\prime}(A+B)\underline{\ell}}} = {}_c\Lambda_2 \cdot {}_c\Lambda_3 $$

$$ = \frac{1 - \dfrac{\underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}}{\underline{\ell}^{*\prime}A\underline{\ell}}}{\dfrac{\underline{\ell}^{*\prime}A\underline{\ell}}{\underline{\ell}^{*\prime}(A+B)\underline{\ell}}} \cdot \frac{\dfrac{|A|}{|A+B|}}{1 - \dfrac{\underline{\ell}^{*\prime}B(A+B)^{-1}B\underline{\ell}}{\underline{\ell}^{*\prime}A\underline{\ell}}} $$

where ${}_c\Lambda_2 \sim \Lambda(n-1,\ p-1,\ 1)$ independently of ${}_c\Lambda_3 \sim \Lambda(n-1,\ p-1,\ q-1)$; where ${}_c\Lambda_2$ is the direction factor and ${}_c\Lambda_3$ is the 'partial' collinearity factor.

CHAPTER IV

SUMMARY

Through the work of Wooding (1956) and then Goodman (1963) the foundation for complex multivariate analysis was established. This work was extended by the major contribution of Khatri (1965) who developed much of the basic theory needed for the analysis of complex multivariate normal random variables. It was this author's intentions of further extending this development by examining the complex analogue of Wilks' statistic as used in the multivariate analysis of variance procedure and as used in the discrimination problem.

Chapter II of this paper contains the basic theorems necessary for the development of the analogue of Wilks' statistic as developed in Chapter III. Also included are results pertaining to the decomposition of a complex Wishart matrix. This decomposition was fundamental in establishing results about the coherence between two complex random variables and the coherence between two complex random vectors. One possible extension of this area is the investigation of the complex t-distribution which occurs as the distribution of the sample estimate when the true coherence between two complex normal variates is zero. This is in direct analogy with the real case.

The development pertaining to Wilks' statistic was presented in Chapter III. In Section 1 of Chapter III, a test for the equality of means of k complex p-variate normal populations was developed. This was accomplished by use of the likelihood ratio criterion. A possible extension of this section would be to examine the properties of this test procedure when applied to complex random variables. Some work has been done in this area by Giri (1965). In Section 2, a test for dimensionality of the mean space was developed and it was shown that this was equivalent to knowing the number of linear discriminant functions needed to discriminate among the k-populations. A needed extension is to determine the exact correction factor that should be used for the chi-square test. In Section 3, a test for the goodness of fit of a single hypothetical discriminant function was developed. The test statistic was factored into two independent parts, one for the direction of the hypothetical discriminant function and the other to test the 'partial' collinearity of the means. A possible extension would be the goodness of fit of more than one discriminant function. Another extension would be the goodness of fit of a single discriminant function from the vector space of dummy variables.

For the p-variate complex normal as defined by Wooding and Goodman, Khatri (1964) has noted "that one can handle all the classical problems of point estimation and testing hypotheses concerning the parameters of complex multivariate normal populations much as one handles those for multivariate populations in real variables." This is true for the problems that have been considered in this paper. As a matter of fact, all the work so far with the p-variate complex normal as defined by Wooding and Goodman has been done by paralleling the real case. Even so, this type of development is important because complex multivariate analysis has possible use and applications in stochastic processes and spectral analysis of multivariate time series and point processes. N. R. Goodman and M. R. Dubman (1968) have considered the theory of time-varying spectral analysis and complex Wishart matrix processes and the assumption of stationarity and normality. They have considered complex Gaussian processes, Wishart processes and time-varying spectral estimates along with distributional results associated with them. This author feels that the work reported in this dissertation will be helpful, indirectly at least, in such investigations.

David Brillinger (1968) has considered the canonical analysis of stationary time series. The distributional results about canonical correlation of complex normal vectors are likely to be useful in this area.

LIST OF REFERENCES

1. Bartlett, M. S. (1951). "Goodness of fit of a single hypothetical discriminant function in the case of several groups," Annals of Eugenics, 16, 199-214.

2. Brillinger, D. R. (1968). "The canonical analysis of stationary time series," Multivariate Analysis II, P. R. Krishnaiah, Ed., Academic Press, New York and London, 331-349.

3. Giri, N. (1965). "On the complex analogues of T^2 and R^2 tests," The Annals of Mathematical Statistics, 36, 664-671.

4. Gupta, A. K. (1971). "Non-null distribution of Wilks' likelihood ratio criterion." (To appear in Annals of Mathematical Statistics.)

5. Goodman, N. R. (1963). "Statistical analysis based on a certain multivariate complex Gaussian distribution," The Annals of Mathematical Statistics, 34, 152-178.

6. ________, and Dubman, M. R. (1968). "Theory of time-varying spectral analysis and complex Wishart matrix processes," Multivariate Analysis II, P. R. Krishnaiah, Ed., Academic Press, New York and London, 351-365.

7. Hodgson, V. (1967). "On sampling distribution of the multiple correlation coefficient." (Presented at IMS meeting in Washington, D.C., December 1967.)

8. Hohn, F. E. (1958). Elementary Matrix Algebra. The MacMillian Company, New York.

9. Kabe, D. G. (1966a). "Complex analogue of some classical non-central multivariate distributions," Australian Journal of Statistics, 8, 99-103.

10. ________. (1966b). "Complex analogues of some results in classical normal multivariate regression theory," Australian Journal of Statistics, 8, 87-91.

11. Khatri, C. G. (1964). "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," Annals of Mathematical Statistics, 35, 1807-1810.

12. ________. (1965). "Statistical analysis based on a certain multivariate complex distribution," Annals of Mathematical Statistics, 34, 98-114.

13. Kshirsagar, A. M. (1959). "Bartlett decomposition and Wishart distribution," Annals of Mathematical Statistics, 30, 238-241.

14. ________. (1969). "Correlation between two vector variables," Journal of Royal Statistical Society, 31, 477-485.

15. ________. (1970). "Direction and collinearity statistics in discriminant analysis." (To appear in Calcutta Statistical Association Bulletin.)

16. ________. (1971). "Degree of freedom for the chi-square test of dimsionality." (To appear in the American Statistician.)

17. Rao, C. R. (1965). Linear Statistical Inference and Its Application. Wiley and Sons, New York.

18. Ruben, H. (1966). "Some new results on the distribution of the sample correlation," Journal of Royal Statistical Society, 31, 477-485.

19. Srivastava, M. S. (1965). "On the complex Wishart distribution," Annals of Mathematical Statistics, 36, 313-315.

20. Wani, J. K. and Kabe, D. G. (1970). "On a certain minimization problem," The American Statistician, 24, 29-30.

21. Wooding, R. A. (1956). "The multivariate distribution of complex normal variates," Biometrika, 43, 212-225.