I give consent to this copy of my thesis, when deposited in the University Library, being available for loan and photocopying.

Date 1st July 1969 Signed ....

BESSEL FUNCTIONS OF MATRIX ARGUMENT WITH STATISTICAL APPLICATIONS

by

BRIAN G. LEACH

B.Sc. (Hons.) (Adel.)

Thesis submitted for the Degree of

Doctor of Philosophy

in the University of Adelaide,

Department of Statistics,

June, 1969.

TABLE OF CONTENTS

SUMMARY

Both the non-central Wishart and non-central means with known covariance distributions can be written as the appropriate central distribution multiplied by a factor which in each case involves a ${}_0F_1$ hypergeometric (or Bessel) function of matrix argument (JAMES [3]). The results of this thesis constitute an assault on the problem of evaluating the Bessel functions via asymptotic expansions or exact series for arbitrary argument matrices.

In the first part of this thesis matrix transformations and group integrations are used on the integral representations for the Bessel functions to reduce them to a form suitable for the application of a method of approximation due to G.A. ANDERSON [1]. Asymptotic expansions are derived and these are shown to be valid for large values of the latent roots of the argument matrix or matrices. For the non-central means with known covariance distribution the expansion is used to compute maximum marginal likelihood estimates for the non-centrality parameters and to establish a modified Chi-square test on the number of non-zero non-centralities.

For the Bessel function of one argument matrix I use a differential equation to derive an approximation asymptotic in the number of degrees of freedom. The result is applied to the likelihood factor of the non-central Wishart.

In the latter part of this thesis I consider methods for the direct evaluation of the Bessel functions in terms of series of zonal polynomials and Laguerre polynomials (CONSTANTINE [2]).

By using the Laplace transform for matrix variables I prove some generalisations of classical summation formulae involving the Laguerre polynomial. A summation formula for the determination of the coefficients $\binom{\kappa}{\nu}$ ($a_{\kappa\tau}$ CONSTANTINE [2]) is proved, as well as other identities involving them. These coefficients are then tabulated for the values k=5,6. Incidentally an algorithm for calculating the $g^{\kappa}_{\nu\mu}$, involved in expressing a product of two zonal polynomials in terms of zonal polynomials, is developed.

JAMES [4] has shown that the zonal polynomials can be expressed in terms of the monomial symmetric functions, where the coefficients are easily determined recursively. I calculate these for the direct evaluation of the Bessel functions in zonal polynomial expansions. By summing the first few terms of the series it is possible to study convergence for various argument matrices.

The final section is devoted to making numerical comparisons of all the methods and giving some idea of their ranges of usefulness.

In appendices I give details of the computer programs used as well as considering problems such as the generation and storage of partitions and the indexing of arrays. ..

REFERENCES

[1] G.A. ANDERSON: An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix, Ann. Math. Statist. 36 (1965), 1153-1173.

[2] A.G. CONSTANTINE: The distribution of Hotelling's generalised T_0^2, Ann. Math. Statistics 37 (1966), 215-225.

[3] A.T. JAMES: The distribution of non central means with known covariance, Ann. Math. Statist. 32 (1961), 874-882.

[4] A.T. JAMES: Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator, Ann. Math. Statist. 39 (1968), 1153-1173.

SIGNED STATEMENT

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University. To the best of my knowledge and belief, the thesis contains no material previously published or written by any other person, except where due reference is made in the text of the thesis.

(B.G. Leach)

ACKNOWLEDGEMENTS

The author is indebted to his supervisor Professor A.T. James for suggesting the topic for research and for encouragement and assistance given throughout this work. He has also benefited from discussions with Dr. A.G. Constantine of the Commonwealth Scientific and Industrial Research Organisation (Division of Mathematical Statistics), who also suggested the use of the system of partial differential equations (Chapter 5).

The author gratefully acknowledges the financial support of a Commonwealth Postgraduate Award and the use of computing facilities at the University of Adelaide. Many thanks are extended to Miss D.J. Potter for her efficient typing of the manuscript, and to Mr. R. Goodwins for the photographic work. Finally the author expresses a vote of thanks to his wife Carolyn for the duplication of the thesis.

CHAPTER 1

INTRODUCTION

1.1 General

The topic is multivariate normal analysis based on the multivariate normal distribution. Let the $m\times n$ matrix variate X, with $m \leqslant n$, be distributed as

$$dF(X;M,\Sigma) = (2\pi)^{-\frac{1}{2}mn}\det \Sigma^{-\frac{1}{2}n}\operatorname{etr}\{-\tfrac{1}{2}\Sigma^{-1}(X-M)(X-M)'\}\prod_{i,j} dx_{ij} \tag{1.1}$$

where $E[X] = M$, $X = (x_1 \ldots x_i \ldots x_n)$, $x_i = (x_{ji})$ $m\times 1$ and $\operatorname{cov}(x_i, x_j) = \delta_{ij}\Sigma$. That is, the columns of X form n independent samples with x_i from $N(m_i,\Sigma)$, where $M = (m_1 \ldots . m_n)$.

In 1961 JAMES [20] has given the non-central Wishart distribution, which is the distribution of

$$XX', \tag{1.2}$$

and the non-central means with known covariance matrix distribution, that is the distribution of the latent roots $w_1, w_2 \ldots w_m$ of the determinantal equation $(w_1 \geqslant w_2 \geqslant \ldots \geqslant w_m)$

$$\det(XX'-w\Sigma) = 0. \tag{1.3}$$

The central distribution of XX' (i.e. $M=0$) was given by WISHART [34] as

$$\frac{1}{2^{\frac{1}{2}mn}\det \Sigma^{\frac{1}{2}n}\Gamma_m(\frac{1}{2}n)} \det(XX')^{\frac{1}{2}(n-m-1)}\operatorname{etr}(-\tfrac{1}{2}\Sigma^{-1}XX')\,(d(XX')) \tag{1.4}$$

where

$$\Gamma_m(a) = \pi^{\frac{1}{4}m(m-1)} \prod_{i=1}^{m} \Gamma(a - \tfrac{1}{2}(i-1)) \tag{1.5}$$

and $\Gamma(a)$ is the ordinary Gamma function. For the latent roots $w_1, w_2, \ldots . w_m$ the central joint distribution was given by FISHER [11], HSU [13] and ROY [31] as

$$\frac{\pi^{\frac{1}{2}m^2}}{2^{\frac{1}{2}mn}\Gamma_m(\frac{1}{2}n)\Gamma_m(\frac{1}{2}m)} \exp(-\tfrac{1}{2}\sum_{i=1}^{m} w_i) \prod_{i=1}^{m} w_i^{\frac{1}{2}(n-m-1)} \prod_{i<j} (w_i - w_j) \prod_{j=1}^{m} dw_i . \tag{1.6}$$

Now if (1.1) is written as

$$dF(X;M,\Sigma) = \mathrm{etr}(-\tfrac{1}{2}\Sigma^{-1}MM')\mathrm{etr}(M'\Sigma^{-1}X)dF(X;0,\Sigma) \tag{1.7}$$

then both non-central distributions can be written as a likelihood factor multiplied by the appropriate central distribution. That is, the non-central distribution of XX' is

$$c_1 \ \mathrm{etr}(-\tfrac{1}{2}\Sigma^{-1}MM') \int_{\mathcal{O}(n)} \mathrm{etr}(M'\Sigma^{-1}XH)(dH)\times(1.4) \tag{1.8}$$

where (dH) stands for the invariant Haar measure on the group $\mathcal{O}(n)$ of $n\times n$ orthogonal matrices H and

$$\mathrm{Vol}(\mathcal{O}(n)) = \int_{\mathcal{O}(n)} (dH) = \frac{2^n \pi^{\frac{1}{2}n^2}}{\Gamma_n(\frac{1}{2}n)} \tag{1.9}$$

making $c_1 = [\mathrm{Vol}(\mathcal{O}(n))]^{-1}$ to give the integral of (1.8) the value unity for $M=0$. The process of integration over $\mathcal{O}(n)$ is called averaging (see JAMES [16],[17]). One further integration gives the non-central distribution of the roots $w_1 \ldots w_m$ as

$$c_2 \operatorname{etr}(-\tfrac{1}{2}\Sigma^{-1}MM') \int_{\mathcal{O}(m)} (dH_1) \int_{\mathcal{O}(n)} (dH_2) \operatorname{etr}((\Sigma^{-\frac{1}{2}}M)' H_1 (\Sigma^{-\frac{1}{2}}X) H_2) \times (1.6) \tag{1.10}$$

with the normalising constant $c_2 = [\mathrm{Vol}(\mathcal{O}(n))\mathrm{Vol}(\mathcal{O}(m))]^{-1}$.

In [20], JAMES also showed how to expand both integrals in series of zonal polynomials. These polynomials $Z_\kappa(S)$, where S is an $m \times m$ symmetric matrix, are homogeneous symmetric polynomials in the latent roots of S corresponding to the partitions

$$\kappa = (k_1\ k_2 \ldots . k_m)\ k_1 \geqslant k_2 \geqslant \ldots . \geqslant k_m \tag{1.11}$$

of the integer k into not more than m parts. A most important property is their average over the orthogonal group $\mathcal{O}(m)$, given by

$$\int_{\mathcal{O}(m)} Z_\kappa(RH'SH)(dH) = \frac{Z_\kappa(R)Z_\kappa(S)}{Z_\kappa(I_m)} \tag{1.12}$$

where I_m is the $m \times m$ identity matrix and

$$Z_\kappa(I_m) = 2^k(\tfrac{1}{2}m)_\kappa \tag{1.13}$$

with

$$(a)_\kappa = \prod_{i=1}^{m} (a-\tfrac{1}{2}(i-1))_{k_i} \quad (a)_n = a(a+1)\ldots(a+n-1). \tag{1.14}$$

Full definitions, proofs etc. can be found in JAMES [18], [19], [20] and CONSTANTINE [7].

Subsequently in 1963, CONSTANTINE [7] gave a power series representation of the hypergeometric functions of matrix argument. The more general form has two argument matrices and is defined as

$$ {}_pF_q^{(m)}(a_1 \ldots a_p; b_1 \ldots b_q; R, S) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_\kappa \ldots (a_p)_\kappa}{(b_1)_\kappa \ldots (b_q)_\kappa} \frac{C_\kappa(R) C_\kappa(S)}{k! \; C_\kappa(I)} \tag{1.15} $$

with the summation over all partitions κ of k, and

$$ C_\kappa(S) = \frac{\chi_{[2\kappa]}(1)}{1.3 \ldots (2k-1)} Z_\kappa(S) \tag{1.16} $$

with $\chi_{[2\kappa]}(1)$ the dimension of the representation $[2\kappa]$ of the symmetric group on 2k symbols. This change in normalisation (1.16) simplifies the formulae. For one argument we have

$$ {}_pF_q(a_1 \ldots a_p; b_1 \ldots b_q; R) = {}_pF_q^{(m)}(a_1 \ldots a_p; b_1 \ldots b_q; R, I) \tag{1.17} $$

and from (1.12)

$$ {}_pF_q^{(m)}(a_1 \ldots a_p; b_1 \ldots b_q; R, S) = \int_{O(m)} {}_pF_q(a_1 \ldots a_p; b_1 \ldots b_q; RH'SH)(dH). \tag{1.18} $$

Full definitions, proofs etc. are found in CONSTANTINE [7] and JAMES [21].

Using the hypergeometric notation JAMES [21] gives (1.8) as

$$ \operatorname{etr}(-\tfrac{1}{2}\Sigma^{-1}MM')\,{}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}\Sigma^{-1}MM'\Sigma^{-1}XX') \times (1.4) \tag{1.19} $$

and if we write $W = \operatorname{diag}(w_i)$, $\Omega = \operatorname{diag}(\omega_i)$ where the ω_i are the roots of

$$\det(MM' - \omega\Sigma) = 0 \tag{1.20}$$

then (1.10) becomes

$$\mathrm{etr}(-\tfrac{1}{2}\Omega)\,{}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W)\times(1.6)\,. \tag{1.21}$$

The ${}_0F_1$ is called the Bessel function of matrix argument and is a generalisation of the familiar univariate function defined by

$${}_0F_1(a;x) = \sum_{n=0}^{\infty} \frac{x^n}{(a)_n n!}\,. \tag{1.22}$$

This function appears in the non-central χ^2. If variates x_i are independent $N(0,1)$ then the distribution of $\chi^2 = (x_1 + \sqrt{\omega})^2 + x_2{}^2 + \ldots + x_n{}^2$ is (setting $w = \chi^2$)

$$e^{-\frac{1}{2}\omega}\,{}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}\omega w)\,\frac{1}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2}n)}\,e^{-\frac{1}{2}w}\,w^{\frac{1}{2}n-1}\,dw. \tag{1.23}$$

Both (1.19) and (1.21) reduce to (1.23) if $m=1$ but (1.21), or the distribution of the non-central means with known covariance, is considered to be the generalisation of the non-central χ^2.

1.2 Historical

The first results on the non-central Wishart distribution were obtained in 1944 by T.W. ANDERSON and GIRSHICK [3]. They stated the general problem in the form of a multiple integral and gave the solution for the rank of $M \leq 2$. Both results are expressed in terms of the Bessel functions $I_\nu(x)$. Subsequently in 1947 ANDERSON [2], by transforming the general multiple integral, managed to perform some of the required integrations and produced an integral representation that appeared to be a matrix analogue of the Poisson integral representation of the Bessel function. The distribution for rank 3 was obtained in 1953 by WEIBULL [33] and in 1955, prior to the zonal expansions, JAMES [16],[17] gave a power series expansion for the general distribution.

The first definite mention of the non-central means with known covariance distribution was in 1961 when JAMES [20] gave the general distribution. It has subsequently been studied in JAMES [21], [22] where it was shown to be the limiting distribution for the general non-central means distribution with finite error degrees of freedom and for the canonical correlations distribution both of which were derived by CONSTANTINE [7].

For convenience these two limiting processes will be outlined here. If X is such that the $m \times m$ matrix XX' has the non-central Wishart distribution on s degrees of freedom with non-centrality parameters Ω (defined by (1.20)) and Y is such that YY' has the central Wishart distribution on t degrees of freedom, the covariance matrix in each case being Σ, then the distribution of the latent roots $r_1 \ldots r_m$ of the matrix $R = XX'(XX'+YY')^{-1}$ is given by

$$\mathrm{etr}(-\tfrac{1}{2}\Omega)\,{}_1F_1^{(m)}(\tfrac{1}{2}(s+t);\tfrac{1}{2}s;\tfrac{1}{2}\Omega,R)$$

$$\frac{\pi^{\frac{1}{2}m^2}\Gamma_m(\frac{1}{2}(s+t))}{\Gamma_m(\frac{1}{2}s)\Gamma_m(\frac{1}{2}t)\Gamma_m(\frac{1}{2}m)} \prod_{i=1}^{m} r_i^{\frac{1}{2}(s-m-1)}(1-r_i)^{\frac{1}{2}(t-m-1)} \prod_{i<j}(r_i-r_j)\Pi dr_i . \tag{1.24}$$

The r_i are the solutions of

$$\det(XX'(XX' + YY')^{-1} - rI) = 0 \tag{1.25}$$

which can be rearranged as

$$\det(XX' - \frac{r}{1-r} YY') = 0. \tag{1.26}$$

Now set $YY' = tS$ (so that as $t \to \infty$, $S \to \Sigma$) then if

$$w = \frac{r}{1-r}\, t \tag{1.27}$$

as $t \to \infty$ (1.26) becomes $\det(XX' - w\Sigma) = 0$ and it is quite easy to show that under the substitution (1.27) the limit of the distribution (1.24) as $t \to \infty$ is (1.21).

Also if $r_1^2 \ldots r_m^2$ are the sample canonical correlation coefficients and $\rho_1^2 \ldots \rho_m^2$ the population canonical correlation coefficients and there are t degrees of freedom, then the limiting distribution, as $t \to \infty$, of

$$tr_i^2 = w_i \quad \text{such that} \quad t\rho_i^2 = \omega_i \tag{1.28}$$

is again (1.21).

The most recent reference to the non-central means with known covariance distribution was in 1968 when PILLAI and GUPTA [28] showed that the first two moments of $W_2^{(m)}$, the second elementary symmetric function in $\frac{1}{2}w_1 \ldots \frac{1}{2}w_m$, could be expressed in terms of generalised Laguerre polynomials. The results are of interest in what follows for the appearance of these polynomials whose evaluation will be considered at length.

1.3 Review of problems considered

By direct summation of the zonal series the Bessel functions can be evaluated for argument matrices with small latent roots. For medium sized latent roots the Bessel function is expressed in terms of the generalised Laguerre polynomials (see CONSTANTINE [8]) in an attempt to obtain more rapidly converging series. The results and details are given in Chapters 6 and 7.

Two Chapters are devoted to the problem of expanding the Bessel function in an asymptotic series valid when the latent roots are large. The starting point in each case is the integral representation i.e. (1.10) and (1.8). The method used is an extension of that used in 1965 by G.A. ANDERSON [1] to expand the hypergeometric function ${}_0F_0^{(m)}(-\frac{1}{2}n\Sigma^{-1},L)$ in an asymptotic series for large n. He begins with the integral definition

$$ {}_0F_0^{(m)}(-\tfrac{1}{2}n\Sigma^{-1},L) = \int_{\mathcal{O}(m)} \operatorname{etr}(-\tfrac{1}{2}n\Sigma^{-1}H'LH)(dH) \tag{1.29} $$

where $H \in \mathcal{O}(m)$ and $L = \operatorname{diag}(\ell_i)$ where the ℓ_i are the latent roots given by $\det(XX' - \ell I_m) = 0$. In Chapter 2 the more difficult two argument case is considered and the results are easily modified in Chapter 4 to deal with the simpler one argument function.

The approximations are a generalisation of the classical result for a ${}_0F_1$ with a large variable. The ${}_0F_1$ has the Poisson integral representation (ERDELYI [9] 7.12(10) and 7.2.2(12))

$$ {}_0F_1(c;x) = \frac{\Gamma(c)}{\Gamma(\frac{1}{2})\Gamma(c-\frac{1}{2})} \int_{-1}^{1} e^{2\sqrt{x}t}(1-t^2)^{c-\frac{3}{2}}dt. \qquad (1.30) $$

The three steps in the approximation process are:-

1. Transform $1-t = u$

$$ = \frac{\Gamma(c)2^{c-\frac{3}{2}}e^{2\sqrt{x}}}{\Gamma(\frac{1}{2})\Gamma(c-\frac{1}{2})} \int_{0}^{2} e^{-2\sqrt{x}u}u^{c-\frac{3}{2}}(1-\frac{u}{2})^{c-\frac{3}{2}}du. $$

2. Expand binomially

$$ = \frac{\Gamma(c)2^{c-\frac{3}{2}}e^{2\sqrt{x}}}{\Gamma(\frac{1}{2})\Gamma(c-\frac{1}{2})} \sum_{r=0}^{\infty} \binom{c-\frac{3}{2}}{r} \frac{(-1)^r}{2^r} \int_{0}^{2} e^{-2\sqrt{x}u}u^{c-\frac{3}{2}+r}du. $$

3. For large x the major contribution to the integral is made near the origin and adding in the tail $2 < u < \infty$ does not alter the value much, thus

$$ \int_{0}^{2} f(u)du \simeq \int_{0}^{\infty} f(u)du $$

and

$$ {}_0F_1(c;x) \simeq \frac{\Gamma(c)e^{2\sqrt{x}}}{2\sqrt{\pi}\, x^{\frac{c}{2}-\frac{1}{4}}} \sum_{r=0}^{\infty} \frac{(-1)^r\,\Gamma(c-\frac{1}{2}+r)}{r!\;\Gamma(c-\frac{1}{2}-r)} \left(\frac{1}{4\sqrt{x}}\right)^r . \qquad (1.31) $$

Both asymptotic results will be shown to agree with this for m=1.

Finally in Chapter 5 a differential equation approach is used to give an asymptotic approximation of ${}_0F_1(\frac{1}{2}n;R)$ valid for large n. This is a generalisation of the one variable method (see e.g. ERDELYI [10]).

Some statistical applications of the non-central means with known covariance distribution are given in Chapter 3. In particular problems of parameter estimation and hypothesis testing by maximum likelihood are considered. The main use of the approximations to the non-central Wishart is to permit the likelihood function to be calculated over a large part of the range. There are difficulties in evaluating the ${}_0F_1$ when both n and the latent roots are both large, but the term $\mathrm{etr}(-\frac{1}{2}\Omega)$ should prevent the likelihood being large in this region.

1.4 Notation

Two further group integrations will be considered:-

1. $$\int_{\mathcal{V}_{nm}} (dH_1) \qquad \mathrm{Vol}(\mathcal{V}_{nm}) = \frac{2^m \pi^{\frac{1}{2}mn}}{\Gamma_m(\frac{1}{2}n)} \tag{1.32}$$

where $H_1 \in \mathcal{V}_{nm}$, the Stiefel manifold, and H_1 consists of the first m columns of the matrix $H \in \mathcal{O}(n)$ i.e.

$$H = [\overset{m}{H_1}\ \overset{n-m}{H_2}]n \qquad H_1'H_1 = I_m \tag{1.33}$$

and (dH_1) stands for the invariant Haar measure on the group $\mathcal{V}_{nm}$.

2. For S $m \times m$ symmetric we will consider integrals of the form

$$\int_{S>0} (dS) \qquad \int_{0<S<T} (dS) \tag{1.34}$$

where the inequality $S>0$ is understood to mean S is positive definite, $S<T$ means $T-S>0$, the integrations are over the space of positive definite matrices and $(dS) = \bigwedge_{i \leqslant j}^{m} ds_{ij}$. $\wedge \equiv$ exterior product. Full definitions and proofs are given in JAMES [15].

The following definitions are from ERDELYI [10].

<u>DEFINITION 1</u>. The sequence of functions $\{\varphi_n(x)\}$ is an <u>asymptotic sequence</u> as $x \to \infty$, if for each n

$$\varphi_{n+1}(x) = o(\varphi_n(x)) \quad \text{as} \quad x \to \infty.$$

Let $\{\varphi_n(x)\}$ be an asymptotic sequence.

<u>DEFINITION 2</u>. The (formal) series $\Sigma\, a_n\, \varphi_n(x)$ is an <u>asymptotic expansion</u> to N terms of $f(x)$ as $x \to \infty$ if

$$f(x) = \sum^{N} a_n\, \varphi_n(x) + o(\varphi_N(x)) \quad \text{as} \quad x \to \infty.$$

This is written $f(n) \sim \Sigma^N a_n\, \varphi_n(x)$.

<u>DEFINITION 3</u>. The function $\varphi(x)$ is an <u>asymptotic representation</u> for $f(x)$ as $x \to \infty$ if

$$f(x) = \varphi(x) + o(\varphi(x)) \quad \text{as} \quad x \to \infty.$$

This is written $f(x) \sim \varphi(x)$.

Let $u_1, u_2 \ldots u_m$ be indeterminates. The elementary symmetric functions a_i are defined by

$$a_1 = \Sigma u_i$$

$$a_2 = \sum_{i<j} u_i u_j$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot$$

$$a_m = u_1 u_2 \ldots u_m .$$

The power sums r_i corresponding to the one part partitions (i) have the definition

$$r_i = \sum_{n=1}^{m} u_n^i .$$

In Chapters 2,3 and 4 we will be dealing with "matrices A with large latent roots" and for convenience this is shortened to "large matrices A". Also note the following convention for the use of the "little-o" notation. Let

$$h_{ij} = 1 - \tfrac{1}{2} \sum_{j=1}^{m} s_{ij}^2 + o(s^2)$$

where $o(s^2)$ means that all subsequent terms are at least third degree in the s_{ij}. Similarly if also

$$k_{ij} = 1 - \tfrac{1}{2} \sum_{j} t_{ij}^2 + o(t^2)$$

then

$$h_{ij}k_{ij} = 1 - \tfrac{1}{2} \sum_{j} (s_{ij}^2 + t_{ij}^2) + o(s^2)$$

where here $o(s^2)$ means that all subsequent terms are at least third degree in the s_{ij}, t_{ij}.

It is also convenient to have a second notation for partitions. Let $\kappa = (k_1 k_2 \ldots k_m) = (1^{\pi_1} 2^{\pi_2} \ldots i^{\pi_i} \ldots)$ where

π_1 of the k_j are 1, π_2 are $2 \ldots \pi_i$ are i and if $k_1 = s$ then $\pi_r = 0$ for $r > s$. An ordering of partitions of a given integer can also be introduced. Let $\kappa = (k_1 k_2 \ldots k_m)$ and $\lambda = (\ell_1 \ell_2 \ldots \ell_m)$ be two partitions of k, then $\kappa > \lambda$ if $k_1 = \ell_1 \ldots k_i = \ell_i, k_{i+1} > \ell_{i+1}$. This will prove useful for specifying the range of summation for summing over sub-sets of partitions. This ordering also leads to the defin-ition of weights. Let $u_1 \ldots u_m$ be indeterminates and $u_1^{k_1} \ldots u_m^{k_m}$, $u_1^{\ell_1} \ldots u_m^{\ell_m}$ be two monomials with indices κ and λ respectively. Then if $\kappa > \lambda$ the former monomial is said to be of <u>higher</u> <u>weight</u>.

The summation conventions to be observed for partit-ions are

1. $\sum_{\kappa} \equiv$ sum over all partitions κ of a particular k and
2. $\sum_{k=0}^{\infty} \sum_{\kappa}$ shortens to $\sum_{k,\kappa}$ where possible.

Finally the LEMMA'S 2.1 and 2.2 of Chapter 2 are both well known but no sources are acknowledged.

CHAPTER 2

THE ASYMPTOTIC FORMULA FOR ${}_0F_1^{(m)}(\frac{1}{2}n; \frac{1}{4}\Omega,W)$

2.1 Introduction

In this chapter an asymptotic expansion for the Bessel function ${}_0F_1^{(m)}(\frac{1}{2}n; \frac{1}{4}\Omega,W)$ is given. The expansion is shown to be valid for matrices Ω and W with large latent roots. A method of approximation developed by G.A. ANDERSON [1] is used.

The starting point is the integral representation (1.10) and in order to make ANDERSON'S methods work, some variables must be integrated out. In the first stages of these integrations a Poisson type integral representation (2.15) for the Bessel function is found. This is a two matrix argument extension of the integral called ANDERSON'S integral by JAMES [21] equation (151).

Some of the further integrations and transformations are shown to be generalisations of those given in section 1.3 in relation to the classical Bessel function. The validity of the asymptotic formula itself is shown to follow from a LEMMA due to HSU [14].

The statistical implications of the result, summarised in THEOREM 2.4 of section 9, are left to the next Chapter. An alternative derivation of the main term of the asymptotic series will also be given in Chapter 3.

2.2 As a function of the latent roots

From (1.21) and (1.10) the integral representation is

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\ \tfrac{1}{4}\Omega,W) = k_1 \int_{\mathcal{O}(m)} (dH_1) \int_{\mathcal{O}(n)} (dH_2)\ \mathrm{etr}[(\Sigma^{-\frac{1}{2}}M)' H_1 \Sigma^{-\frac{1}{2}} X H_2] \tag{2.1} $$

where $\qquad k_1 = [\mathrm{Vol}(\mathcal{O}(m))\mathrm{Vol}(\mathcal{O}(n))]^{-1}.$

Make the substitutions

$$ R = \Sigma^{-\frac{1}{2}}M\,,\ \det(MM' - \omega\Sigma) = \det(RR' - \omega I) = 0 $$

$$ S = \Sigma^{-\frac{1}{2}}X\,,\ \det(XX' - w\Sigma) = \det(SS' - wI) = 0 $$

and (2.1) becomes

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\ \tfrac{1}{4}\Omega,W) = k_1 \int_{\mathcal{O}(m)} (dH_1) \int_{\mathcal{O}(n)} (dH_2)\ \mathrm{etr}(R'H_1SH_2)\ . \tag{2.2} $$

It is well known that (2.2) is a function of the latent roots ω_i and w_i. This will be demonstrated by applying transformations to R and S that reduce them to a "diagonal" form for rectangular matrices. The elements on the diagonals are simple functions of ω_i and w_i.

We use the well known

LEMMA 2.1

Let Z be a real $m{\times}n$ matrix. Let $A = \mathrm{diag}(a_i)$ where $a_i = +\sqrt{b_i}$ and the b_i are the latent roots of $\det(ZZ' - bI) = 0$. Also $D = (\delta_1\delta_2\ldots\delta_m)$ is the matrix of latent vectors δ_i for ZZ', normalised so that the first element of each column is positive,making D unique. If $E^1 = A^{-1}D'Z$ i.e. $E^1E^{1\prime} = I_m$ then

$$Z = DAE^1. \qquad (2.3)$$

The m×n matrix E^1 forms the first m rows of an orthogonal matrix and we can choose the remaining n-m rows to form the n×n matrix E, where

$$E = \begin{bmatrix} E^1 \\ E^2 \end{bmatrix} \in \mathcal{O}(n). \qquad (2.4)$$

Write (2.3) as

$$Z = DAE^1 = D[A\ 0]\begin{bmatrix} E^1 \\ E^2 \end{bmatrix} = D[A\ 0]E \qquad (2.5)$$

where [A 0] is m×n.

Applying (2.5) to R and S

$$R = D_1[A\ 0]E_1 \qquad D_1, D_2 \in \mathcal{O}(m)$$

where

$$S = D_2[B\ 0]E_2 \qquad E_1, E_2 \in \mathcal{O}(n)$$

and $A = \text{diag}(a_i)$ from $\det(RR' - a^2I) = 0$

$B = \text{diag}(b_i)$ from $\det(SS' - b^2I) = 0$

where the elements of A and B satisfy the ordering

$$a_1 > a_2 > \ldots. > a_m > 0$$

$$b_1 > b_2 > \ldots. > b_m > 0.$$

By comparison with Ω and W, $\omega_i = a_i^2$ and $w_i = b_i^2$.

Substitute in the exponential term of (2.2)

$$\operatorname{etr}(R'H_1SH_2) = \operatorname{etr}\left(\begin{bmatrix} A \\ 0 \end{bmatrix} D_1'H_2D_2[B\ \ 0]E_2H_2E_1'\right)$$

and change variables

$$H_1 \to \underline{H}_1 = D_1'H_1D_2$$

$$H_2 \to \underline{H}_2 = E_2\ H_2E_1'.$$

Since D_1, D_2, E_1, E_2 are constant matrices $(dH_1) = (d\underline{H}_1)$, $(dH_2) = (d\underline{H}_2)$ and on dropping the _ (2.2) becomes

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\ \tfrac{1}{4}\Omega, W) = k_1 \int_{\mathcal{O}(m)} (dH_1) \int_{\mathcal{O}(n)} (dH_2)\ \operatorname{etr}\left(\begin{bmatrix} AH_1B & 0 \\ 0 & 0 \end{bmatrix} H_2\right). \tag{2.6}$$

At this stage the Bessel function is clearly seen to depend only on simple functions $\sqrt{\omega_i}$ and $\sqrt{w_i}$ of the latent roots.

2.3 ANDERSON'S method

G.A. ANDERSON'S [1] approach to approximating integrals of this type is to apply the parameterisation $H = \exp(S)$ where $H \in \mathcal{O}(m)$ and $S = (s_{ij})$ is an $m \times m$ skew symmetric matrix. The s_{ij} are then integrated over the range $(-\infty, \infty)$.

Let us apply this directly to (2.6). Set $H_1 = (h_{ij})$, $H_2 = (k_{ij})$, $S = (s_{ij})$, $T = (t_{ij})$, $A = (a_i \delta_{ij})$ and $B = (b_i \delta_{ij})$. Then

$$H_1 = \exp(S) = I + S + \tfrac{1}{2}S^2 + \cdots$$

$$H_2 = \exp(T) = I + T + \tfrac{1}{2}T^2 + \cdots$$

and the Jacobian is of the form $1 + o(s)$. Element by element

$$\begin{aligned} h_{ii} &= 1 - \tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2 + \cdots & i &= 1,\ldots m \\ h_{ij} &= \quad s_{ij} + \cdots & i &\neq j \\ k_{ii} &= 1 - \tfrac{1}{2}\sum_{j=1}^{n} t_{ij}^2 + \cdots & i &= 1,\ldots n \\ k_{ij} &= \quad t_{ij} + \cdots & i &\neq j. \end{aligned}$$

Now

$$\begin{aligned} \operatorname{tr}\left(\begin{bmatrix} AH_1B & 0 \\ 0 & 0 \end{bmatrix} H_2\right) &= \sum_{i,j,u,v} a_i \delta_{iu} h_{uj} b_j \delta_{jv} k_{vi} \\ &= \sum_{i,j}^{m} a_i b_j h_{ij} k_{ji} \\ &= \sum_{i=1}^{m} a_i b_i h_{ii} k_{ii} + \sum_{i \neq j} a_i b_j h_{ij} k_{ji} \qquad (2.7) \end{aligned}$$

and on substituting, noting that $s_{ij} = s_{ji}$, $s_{ii} = 0$ etc.

$$= \sum_{i=1}^{m} a_i b_i \left(1 - \tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2\right)\left(1 - \tfrac{1}{2}\sum_{j=1}^{n} t_{ij}^2\right) - \sum_{i,j} a_i b_j s_{ij} t_{ij} + \cdots . \qquad (2.8)$$

Now (2.8) does not contain the elements t_{ij}, $i,j = m+1,\ldots n$ so on substitution in (2.6) and extension of the range we have integrals of the form

$$\int_{-\infty}^{\infty} dt_{ij} \qquad i,j = m+1,\ldots n$$

which clearly leads to an absurd result. This is because the variables k_{ij} $i,j = m+1,\ldots n$ are not explicitly involved in the integrand.

We must now try to integrate out those variables only implicitly involved in (2.6) before applying the parameterisation. The elements only involved in the range are

$k_{ij}, k_{ji} \quad i = 1,\ldots m \quad j = m+1,\ldots n$

$k_{ij} \quad i,j = m+1,\ldots n.$

Partitioning H_2 as

$$H_2 = \underset{n}{\begin{bmatrix} H_2^1 \\ H_2^2 \end{bmatrix}} \begin{matrix} m \\ n-m \end{matrix}$$

the integrand of (2.6) reduces to $\operatorname{etr}\left(\begin{bmatrix} AH_1BH_2^1 \\ 0 \end{bmatrix}\right)$. Only elements of H_2^1 appear in the integrand so we integrate over H_2^2 for fixed H_2^1 by the formula

$$\int_{H_2^2} (dH) = k_2(dH_2^1) \qquad (2.9)$$

where (dH_2^1) stands for the invariant volume element on the Stiefel manifold $\mathcal{V}_{nm}$ and

$$k_2 = \frac{\operatorname{Vol}(\mathcal{O}(n))}{\operatorname{Vol}(\mathcal{V}_{nm})}.$$

(See JAMES [15] equations (4.39), (5.10) and (5.16).)

The integral (2.6) becomes

$${}_0F_1^{(m)}(\tfrac{1}{2}n; \tfrac{1}{4}\Omega, W) = k_3 \int_{\mathcal{O}(m)} (dH_1) \int_{\mathcal{V}_{nm}} (dH_2^1)\operatorname{etr}\left(\begin{bmatrix} AH_1BH_2^1 \\ 0 \end{bmatrix}\right) \qquad (2.10)$$

where $\quad k_3 = [\operatorname{Vol}(\mathcal{O}(m))\operatorname{Vol}(\mathcal{V}_{nm})]^{-1}$. Thus we have integrated out k_{ij} $i = m+1,\dots n$, $j = 1,\dots n$.

There are two paths of development to be followed here. Either the Stiefel manifold V_{nm} can be parameterised or the remaining implicit k_{ij} integrated out. In Chapter 3 the parameterisation of V_{nm} is used. The integrating out of the k_{ij} $i = 1,\dots n$, $j = m+1\dots n$ is considered in the next section.

2.4 Reduction to a Poisson type integral

Partition $H_{\frac{1}{2}}$ into

$$H_{\frac{1}{2}} = \underset{m\quad\; n-m}{[H_{\frac{1}{2}}^{1} \quad H_{\frac{1}{2}}^{2}]}\, m.$$

Thus

$$\mathrm{etr}\left(\begin{bmatrix} AH_1BH_{\frac{1}{2}} \\ 0 \end{bmatrix}\right) = \mathrm{etr}\left(\begin{bmatrix} AH_1BH_{\frac{1}{2}}^{1} & AH_1BH_{\frac{1}{2}}^{2} \\ 0 & 0 \end{bmatrix}\right)$$

$$= \mathrm{etr}(AH_1BH_{\frac{1}{2}}^{1})$$

and the integrand now only involves the elements of $H_{\frac{1}{2}}^{1}$.

One more integration can be performed and this is facilitated by using the transformation (HERZ [12], LEMMA 3.7)

$$H_{\frac{1}{2}} = [H_{\frac{1}{2}}^{1}\ H_{\frac{1}{2}}^{2}] = [T \quad (I_m - TT')^{\frac{1}{2}}U']$$

where $H_{\frac{1}{2}}^{1} = T$ real with $TT' \leqslant I_m$, $(I_m - TT')^{\frac{1}{2}}$ is the non-negative square root and $U \in V_{n-m,m}$. The LEMMA imposes the restriction that $n \geqslant 2m$ (against the initial assumption of $n \geqslant m$). This should create no difficulties in a statistical application.

HERZ gives for the measures

$$(dH_2^1) = \det(I-TT')^{\frac{1}{2}(n-2m-1)}(dT)(dU)$$

and the integral (2.10) becomes

$$k_3\int_{\mathcal{O}(m)}(dH_1)\int_{TT'\leqslant I}(dT)\ \mathrm{etr}(AH_1BT)\ \det(I-TT')^{\frac{1}{2}(n-2m-1)}\int_{\mathcal{V}_{n-m,m}}(dU).$$

Integrating U over $\mathcal{V}_{n-m,m}$ gives $\mathrm{Vol}(\mathcal{V}_{n-m,m})$ and

$${}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W)=k_4\int_{\mathcal{O}(m)}(dH_1)\int_{TT'\leqslant I}(dT)\ \mathrm{etr}(AH_1BT)\ \det(I-TT')^{\frac{1}{2}(n-2m-1)} \tag{2.11}$$

where

$$k_4 = \frac{\mathrm{Vol}(\mathcal{V}_{n-m,m})}{\mathrm{Vol}(\mathcal{O}(m))\mathrm{Vol}(\mathcal{V}_{nm})} = \frac{\Gamma_m(\frac{1}{2}m)\Gamma_m(\frac{1}{2}n)}{2^m\pi^{m^2}\Gamma_m(\frac{1}{2}(n-m))}.$$

This integral is the result of averaging ANDERSON'S integral (as quoted by JAMES [21] equation (151)). The derivati on of this integral and the averaging process are given in Chapter 4.

The idea now is to replace the integration over $TT' \leqslant I$ by one or more integrations, one of which is an integration over $\mathcal{O}(m)$.

For T real $m\times m$ we can find an $H_2 \in \mathcal{O}(m)$ such that $T = SH_2$, where $S = (TT')^{\frac{1}{2}}$. That is, S is the symmetric positive semi-definite square root.

To find the Jacobian of the transformation, by (2.3)

$$T = \underline{H}_1' \, D \, \underline{H}_2 \qquad S = \underline{H}_1' \, D \, \underline{H}_1$$

where $\underline{H}_1, \underline{H}_2 \in \mathcal{O}(m)$ and $D = \mathrm{diag}(d_i)$ with the d_i the latent roots of $\det(S - dI) = 0$. Clearly

$$\begin{aligned} T &= \underline{H}_1' \, D \, \underline{H}_1 \quad \underline{H}_1'\underline{H}_2 \\ &= \quad S \qquad\quad H_2 \end{aligned}$$

where $H_2 = \underline{H}_1'\underline{H}_2 \in \mathcal{O}(m)$. Now in [15] if T is a real $m{\times}n$ matrix, then equation (8.8) gives

$$(dT) = (d\underline{H}_1)(d\underline{H}_2) \prod_{i<j} (d_i^2 - d_j^2) \prod_{i=1}^{m} dd_i \tag{2.12}$$

and from (8.19), with $S = (X'X)$ an $m{\times}m$ symmetric matrix,

$$(dS) = (d\underline{H}_1) \prod_{i<j} (d_i - d_j) \prod dd_i . \tag{2.13}$$

Combining (2.12) and (2.13)

$$(dT) = (d\underline{H}_2) \prod_{i<j} (d_i + d_j)(dS). \tag{2.14}$$

But $H_2 = \underline{H}_1'\underline{H}_2$ and for fixed $\underline{H}_1$, as $\underline{H}_2$ ranges over $\mathcal{O}(m)$ so does H_2. This corresponds to having fixed S and as H_2 (or $\underline{H}_2$) ranges over $\mathcal{O}(m)$, T ranges over all matrices for which $TT' = S^2$. Thus when $\underline{H}_1$ is fixed, $(dH_2) = (d\underline{H}_2)$. Substituting in (2.14)

$$(dT) = (dH_2) \prod_{i<j} (d_i + d_j) \, (dS).$$

After the transformation the range of integration is changed from $0 \leqslant TT' \leqslant I$ to $0 \leqslant S^2 \leqslant I$ and since S is assumed symmetric positive semi-definite, this range is equivalently $0 \leqslant S \leqslant I$. The integral (2.11) becomes

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) = k_4 \int_{O(m)} (dH_1) \int_{O(m)} (dH_2) $$

$$ \int_{0\leqslant S\leqslant I} (dS)\,\mathrm{etr}(AH_1BSH_2)\det(I-S^2)^{\frac{1}{2}(n-2m-1)} \prod_{i<j} (d_i + d_j). \tag{2.15} $$

If we set $m=1$ in (2.15) it reduces to (1.30) provided $O(1)$ is considered to consist of the two points ± 1 and the measure (dH) is taken to be 1 at each point (i.e. $\mathrm{Vol}(O(1)) = 2$). Thus this integral can be considered as the two matrix argument analogue of the Poisson integral for the classical Bessel function. We can apply the matrix analogues of steps 1.-3. of section 1.3 to the integral (2.15).

2.5 The classical approximation generalised

Let us first make the substitution as in step 1.

$$ U = I - S \qquad\qquad (\text{cf. } u=1-t). $$

Clearly U is symmetric positive semi-definite and has latent roots $u_i = 1 - d_i$, while the range of integration becomes $0 \leqslant u \leqslant I$. Also it can be shown that

$$ (dU) = (-1)^{\frac{1}{2}m(m+1)}(dS) $$

but since the integrand of (2.15) is always non-negative the sign of the Jacobian can be ignored. Thus the integral becomes

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\ \tfrac{1}{4}\Omega,W) = k_5 \int_{O(m)} (dH_1) \int_{O(m)} (dH_2)\ \mathrm{etr}(AH_1BH_2) \qquad (2.16a) $$

$$ \times \int_{0\leqslant U\leqslant I} (dU)\,\mathrm{etr}(-AH_1BUH_2)[\det U(I-\tfrac{1}{2}U)]^{\frac{1}{2}(n-2m-1)} \prod_{i<j} (1 - \frac{u_i+u_j}{2}) $$

(2.16b)

where $k_5 = 2^{\frac{1}{2}m(n-m-2)} k_4$.

Now following step 2. let us expand

$$ \det(I - \tfrac{1}{2}U)^{\frac{1}{2}(n-2m-1)} $$

in its binomial expansion (CONSTANTINE [7] equation (31)). Also as $\prod_{i<j} (1 - \frac{u_i+u_j}{2})$ is a symmetric function it too has an expansion in zonal polynomials. The two series are then multiplied and the result is expressed in zonal polynomials as

$$ \det(I - \tfrac{1}{2}U)^{\frac{1}{2}(n-2m-1)} \prod_{i<j} (1 - \frac{u_i+u_j}{2}) = 1 + \sum_{k=1}^{\infty} \sum_{\kappa} c_\kappa\ C_\kappa(U). $$

(2.17)

Thus the integral (2.16b) can be written as

$$ \int_{0\leqslant U\leqslant I} (dU)\,\mathrm{etr}(-RU)\det U^{\frac{1}{2}(n-2m-1)} (1 + \sum_{k,\kappa} c_\kappa\ C_\kappa(U)) \qquad (2.18) $$

where $R = H_2AH_1B$.

Finally the analogue of step 3. We want an asymptotic expansion valid for large A and B, so R will be large and $\mathrm{etr}(-RU)$ will be negligible for large values of U. Thus we change the range to $U > 0 (f(U) = 0$ at $U = 0)$

and integrate using a THEOREM of [7].

THEOREM 2.1

Let R be an $m\times m$ positive definite symmetric matrix. Then

$$\int_{S>0} \operatorname{etr}(-RS)\det S^{t-\frac{1}{2}(m+1)} C_\kappa(S)(dS) = \Gamma_m(t,\kappa)\det R^{-t} C_\kappa(R^{-1}). \tag{2.19}$$

The integration is over all positive definite $m\times m$ matrices S and valid for all real t with $t > \frac{1}{2}(m-1)$. Also $\Gamma_m(t,\kappa) = (t)_\kappa \Gamma_m(t)$.

Applying (2.19) term by term to (2.18) with the range of integration extended gives the general term

$$\int_{U>0} \operatorname{etr}(-RU)\det U^{\frac{1}{2}(n-m)-\frac{1}{2}(m+1)} c_\kappa\ C_\kappa(U)(dU)$$

$$= \Gamma_m(\tfrac{1}{2}(n-m))\det R^{-\frac{1}{2}(n-m)} (\tfrac{1}{2}(n-m))_\kappa\ c_\kappa\ C_\kappa(R^{-1})$$

while $\det R = \det(AB)$, $R^{-1} = B^{-1}H_1{}'A^{-1}H_2{}'$. Finally we are left with the evaluation of

$$k_6 \det(AB)^{-\frac{1}{2}(n-m)} \int_{O(m)}(dH_1)\int_{O(m)}(dH_2)\operatorname{etr}(AH_1BH_2)(1+\sum_{k,\kappa} d_\kappa C_\kappa(B^{-1}H_1'A^{-1}H_2')) \tag{2.20}$$

where $k_6 = \Gamma_m(\frac{1}{2}(n-m))k_5$, $d_\kappa = (\frac{1}{2}(n-m))_\kappa c_\kappa$.

This approximate integral representation for ${}_0F_1^{(m)}(\frac{1}{2}n; \frac{1}{4}\Omega,W)$ is now ready to be tackled by parameterisation of H_1 and H_2.

2.6 Finding the maxima

There are two features of ANDERSON'S method:

1. the parameterisation $H = \exp(S)$ is only valid for proper orthogonal matrices H (i.e. those with $\det H = +1$), and
2. the approximation is made in neighbourhoods of the equal maxima of the integrand of (2.20), which is denoted by $f(H_1, H_2; A, B)$ for convenience.

That is we must first convert the integrations of (2.20) into ones over proper orthogonal matrices and then determine the maxima of the integrands.

Firstly $\mathcal{O}(m)$ is the union of two disjoint subsets $\mathcal{O}^+(m)$ and $\mathcal{O}^-(m)$ defined by

$$\mathcal{O}^+(m) = \{H : H \in \mathcal{O}(m), \det H = +1\}$$
$$\mathcal{O}^-(m) = \{H : H \in \mathcal{O}(m), \det H = -1\}.$$

The range of integration of (2.20) breaks into four disjoint ranges, giving

$$\int_{\mathcal{O}(m)}\int_{\mathcal{O}(m)} = \int_{\mathcal{O}^+(m)}\int_{\mathcal{O}^+(m)} + \int_{\mathcal{O}^+(m)}\int_{\mathcal{O}^-(m)} + \int_{\mathcal{O}^-(m)}\int_{\mathcal{O}^+(m)} + \int_{\mathcal{O}^-(m)}\int_{\mathcal{O}^-(m)} \tag{2.21a}$$

or

$$I = I_1 + I_2 + I_3 + I_4. \tag{2.21b}$$

Now it is well known that the elements of $\mathcal{O}^+(m)$ can be mapped one-to-one into those of $\mathcal{O}^-(m)$, and vice versa, by the device of selecting a matrix, A say, from one subset, $\mathcal{O}^-(m)$ say, and then as H runs over all elements of $\mathcal{O}^+(m)$, AH runs over all elements of $\mathcal{O}^-(m)$. In particular, let us consider the mapping

$$H^+ \to H = JH^+ \quad \text{where} \quad J = \begin{bmatrix} I_{m-1} & 0 \\ 0' & -1 \end{bmatrix}.$$

Then $H^+ \in \mathcal{O}^+(m)$ implies that $H \in \mathcal{O}^-(m)$ and since J is fixed $(dH) = (dH^+)$. Applying the transformation to an integral gives

$$\int_{\mathcal{O}^-(m)} y(H)(dH) = \int_{\mathcal{O}^+(m)} y(JH^+)(dH^+).$$

Now returning to (2.21), in I_4 make the transformations

$$H_1 = JH_1^+ \qquad (dH_1) = (dH_1^+)$$
$$H_2 = H_2^+J \qquad (dH_2) = (dH_2^+)$$

and by noting that $JAJ = A$, $JA^{-1}J = A^{-1}$ it is seen that $I_4 = I_1$. Making the same transformations where required in I_2 and I_3 respectively, we obtain (on dropping the $^+$)

$$I_2=I_3= \int_{\mathcal{O}^+(m)} (dH_1) \int_{\mathcal{O}^+(m)} (dH_2) \operatorname{etr}(A^*H_1BH_2)(1+\sum_{k,\kappa} d_\kappa C_\kappa(B^{-1}H_1'A^{*-1}H_2'))$$

where $A^* = AJ = JA = \operatorname{diag}(a_1 \ldots a_{m-1}, -a_m)$. Thus (2.21b) is

$$I = 2I_1 + 2I_2 \tag{2.22}$$

where all integrations are over $O^+(m)$ and we can parameterise $H_1, H_2 \in O^+(m)$.

Secondly we must find the H_1, H_2 that are solutions to

$$\max_{H_1, H_2 \in O^+(m)} f(H_1, H_2; A, B) \tag{2.23}$$

$$\max_{H_1, H_2 \in O^+(m)} f(H_1, H_2; A^*, B). \tag{2.24}$$

Of course it is easier to find all stationary values of $f(H_1, H_2; A, B)$ for $H_1, H_2 \in O(m)$ and identify the two cases afterwards. In fact we restrict ourselves to finding the stationary points of $\text{etr}(AH_1BH_2)$ as this is the dominant factor for large A, B.

Now $\exp(\cdot)$ is a monotone function so we begin by finding all stationary points of $\text{tr}(XHYK')$ over $H, K \in O(m)$, where X, Y are defined by

$$X = \text{diag}(x_i) \qquad x_1 > x_2 > \dots > x_m > 0$$

$$Y = \text{diag}(y_i) \qquad y_1 > y_2 > \dots > y_m > 0.$$

All stationary values of $\text{tr}(XHYK')$ are given by

THEOREM 2.2

If X, Y, H, K are as defined above then the stationary values of the function $\text{tr}(XHYK')$ occur at those points where H, K are signed permutation matrices and related by

$K = I^*H$

$$I^* = \begin{bmatrix} \pm 1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \pm 1 \end{bmatrix}.$$

<u>DEFINITION</u>. Let σ be a permutation of $1,2,\ldots m$. Then by a <u>signed permutation matrix</u> is meant a matrix $P = (p_{ij})$ with

$$p_{ij} = \begin{cases} \pm 1 & i = \sigma(j) \\ 0 & \text{otherwise.} \end{cases}$$

For example $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$

$\sigma(1) = 2$ $\sigma(2) = 3$ $\sigma(3) = 1$

$p_{21} = 1$ $p_{32} = -1$ $p_{13} = 1$

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}.$$

Clearly P is orthogonal.

The proof of the THEOREM uses the well known

LEMMA 2.2

If A is $m \times m$ real and S_α, $\alpha = 1,2,\ldots\frac{1}{2}m(m-1)$ are $m \times m$ skew symmetric and linearly independent, then

$$tr(AS_\alpha) = 0 \qquad \alpha = 1,2,\ldots\tfrac{1}{2}m(m-1)$$

implies that A is symmetric.

<u>Proof</u>

Let $A = (a_{ij})$, $S = (s_{ij})$, then

$$tr(AS) = \sum_{i,j} a_{ij}\, s_{ji} = 0.$$

Rearranging

$$\sum_i a_{ii}\, s_{ii} + \sum_{i<j} (a_{ij} - a_{ji}) s_{ji} = 0. \tag{2.25}$$

Now $s_{ii} = 0 \quad i = 1,\ldots m$ so the a_{ii} are arbitrary. Substituting the elements of $S_1, S_2, \ldots$ in (2.25) gives $\frac{1}{2}m(m-1)$ linearly independent equations in $\frac{1}{2}m(m-1)$ unknowns. The system has the solution $a_{ij} = a_{ji} \; i < j$.

Q.E.D.

<u>Proof</u> (of THEOREM 2.2)

At a stationary point

$$
\begin{aligned}
0 &= d\ \mathrm{tr}(XHYK') \\
&= \mathrm{tr}(XdHYK' + XHYdK') \\
&= \mathrm{tr}(XHH'dHYK' + XHYdK'KK').
\end{aligned}
$$

Now $K'dK$ is skew symmetric ($K'K = I$ implies $dK'K=-K'dK$) as is $H'dH$. Make the substitutions $d\Lambda_1 = H'dH$, $d\Lambda_2 = K'dK$, then

$$\mathrm{tr}(YK'XHd\Lambda_1 - K'XHYd\Lambda_2) = 0. \tag{2.26}$$

Since H,K vary independently over $\mathcal{O}(m)$, then $d\Lambda_1$, $d\Lambda_2$ vary independently over sets of skew matrices and we equate the two terms of (2.26) to zero separately, giving

$$\mathrm{tr}(YK'XHd\Lambda_1) = 0 \quad \mathrm{tr}(K'XHYd\Lambda_2) = 0.$$

Applying the LEMMA 2.2 and putting $C = K'XH$ gives YC and CY symmetric. Then

$$
\begin{aligned}
YC &= C'Y \qquad (2.27)\\
YC' &= CY
\end{aligned}
$$

and adding

$$Y(C+C') = (C+C')Y.$$

Now Y is diagonal so in order to commute with Y we must have $C+C'$ diagonal. Put $C+C' = 2D$, then $C = D+S$ where S is skew symmetric.

Substituting in (2.27)

$$Y(D+S) = (D-S)Y$$

i.e. $$YS + SY = 0.$$

Put $S = (s_{ij})$ and write out element by element as

$$(y_i + y_j)s_{ij} = 0 \qquad i,j = 1,\ldots m.$$

Then $s_{ij} = 0$ $i,j = 1,\ldots m$. Hence C is diagonal. So $K'XH = D$ where $D = \mathrm{diag}(d_i)$.

Now

$$D^2 = D'D = H'X^2H \tag{2.28}$$

$$= DD' = K'X^2K \tag{2.29}$$

and equating the right hand sides of (2.28) and (2.29)

$$X^2KH' = KH'X^2.$$

Put $T = (t_{ij}) = KH'$ and write out element by element as

$$x_i{}^2 t_{ij} = t_{ij} x_j{}^2.$$

Thus when $i \neq j$, $x_i{}^2 \neq x_j{}^2$ making $t_{ij} = 0$ and when $i=j$ $t_{ii} = \pm 1$ since $T \in \mathcal{O}(m)$. Hence $KH' = I^*$ and $K = I^*H$.

Now $D = K'XH = H'I^*XH = H'X^*H$ where $X^* = \mathrm{diag}(\pm x_i)$ and the elements of D are the same as those of X^* but in some rearrangement i.e. $d_i = \pm x_{\sigma(i)}$. From

$$X^*H = HD, \quad x_i h_{ij} = h_{ij}\, x_{\sigma(j)} \tag{2.30}$$

comes

$$h_{ij} = \begin{cases} \pm 1 & i = \sigma(j) \quad \text{i.e. } h_{\sigma(j)j} = \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

which is the definition of a signed permutation matrix H. Also $K = I^*H$ is a signed permutation matrix with $k_{\sigma(j)j} = \pm h_{\sigma(j)j}$ depending on the signs of the diagonal elements of I^*. Q.E.D.

We must now find the stationary points of $\text{etr}(AH_1BH_2)$ and $\text{etr}(A^*H_1BH_2)$ for $H_1, H_2 \in \mathcal{O}^+(m)$. From the above proof we see that at a stationary point the function takes the value

$$\text{tr}(DY) = \sum_{i=1}^{m} \pm x_{\sigma(i)} y_i . \tag{2.31}$$

Clearly the absolute maximum is $\Sigma x_i y_i$ for $i = \sigma(i)$, $d_i = x_i$ and $h_{ii} = \pm 1$. Now $D = X$ and substituting in (2.30) $X^*H = HX$ giving $X^* = X = I^*X$ and $I^* = I$. Thus

$$H = K = \begin{bmatrix} \pm 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & \pm 1 \end{bmatrix} . \tag{2.32}$$

Applying this result to $\text{etr}(AH_1BH_2)$ we have 2^m equal maxima of $\text{etr}(AB)$ at points (H_1, H_2) where $H_1 = H_2 = H$ (as in (2.32)) and of these 2^{m-1} have

$H_1, H_2 \in \mathcal{O}^+(m)$. Thus we will take our asymptotic expansion in the region of those $H_1, H_2 \in \mathcal{O}^+(m)$ for which $\mathrm{etr}(AH_1BH_2)$ is near its maximum of $\mathrm{etr}(AB)$, on the assumption that for large A,B this region will also contain all those H_1, H_2 for which $f(H_1,H_2;A,B)$ is also large. Note also that all the maxima occur at points in the range of integration of I_1 so we can now approximate it.

Now we find the maximum for the dominant term of the integrand of I_2. That is we must find

$$\max_{H_1, H_2 \in \mathcal{O}^+(m)} \mathrm{tr}(A^* H_1 B H_2) \tag{2.33}$$

or in other words

$$\max_{\substack{H_1, H_2 \text{ in different} \\ \text{subsets}}} \mathrm{tr}(AH_1BH_2).$$

From the THEOREM 2.2 $H_2 = H_1' I^*$ $(K = I^*H)$ and for H_1, H_2 to belong to different subsets, I^* must have an odd number of -1's on the diagonal. Let $D = H_2AH_1 = H_1'(I^*A)H_1$, then an odd number of the $d_i = -a_{\sigma(i)}$. Since there must be at least one negative a_i it is clear that

$$\mathrm{tr}(DB) = \sum_{i=1}^{m-1} a_i b_i - a_m b_m$$

where $D = \mathrm{diag}(a_1 \ldots . a_{m-1}, -a_m)$, is the maximum value of $\mathrm{tr}(AH_1BH_2)$ when H_1, H_2 come from different subsets. At this maximum

$$H_1, H_2 = \begin{bmatrix} \pm 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ & & & \pm 1 \end{bmatrix} \quad \text{but} \quad H_2 = H_1 J.$$

There are 2^m solution points in all and 2^{m-1} are solutions to (2.33) and hence are maxima in the range of I_2.

We will see that for large values of A and B the integrands of (2.22) will be negligible apart from small neighbourhoods about each maximum. The integral I_1 will consist of identical contributions from each of these 2^{m-1} neighbourhoods, and similarly for I_2. One solution to the maximisation problem in each case is $H_1 = H_2 = I$ so (2.22) becomes

$$I \simeq 2^m \int_{N(I)} \int_{N(I)} f(H_1, H_2; A, B)(dH_1)(dH_2) \qquad (2.34)$$

$$+ 2^m \int_{N(I)} \int_{N(I)} f(H_1, H_2; A^*, B)(dH_1)(dH_2)$$

where $N(I)$ is a neighbourhood of I on the orthogonal manifold $\mathcal{O}(m)$ and contains only matrices in $\mathcal{O}^+(m)$.

We will now focus attention on the approximation of

$$g(A,B) = \int_{N(I)} \int_{N(I)} f(H_1, H_2; A, B)(dH_1)(dH_2) \qquad (2.35)$$

but the same methods are valid for $g(A^*, B)$.

2.7 Approximating the integral

Now in (2.35) both $H_1, H_2 \in \mathcal{O}^+(m)$ so we apply the parameterisation of section 3. Let $S = (s_{ij})$, $T = (t_{ij})$ be $m \times m$ skew symmetric matrices and $H_1 = (h_{ij})$, $H_2 = (k_{ij})$. Then

$$H_1 = \exp(S) \qquad H_2 = \exp(T)$$

and element by element

$$h_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2 + o(s^2), \quad k_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{m} t_{ij}^2 + o(t^2) \qquad i = 1,2,,,m \quad (2.36a)$$

$$h_{ij} = s_{ij} + o(s), \quad k_{ij} = t_{ij} + o(t) \qquad i \neq j. \quad (2.36b)$$

For the transformation $H_1 = \exp(S)$ the Jacobian is

$$J(S) = 1 + \frac{m-2}{24}\operatorname{tr} S^2 + \frac{8-m}{4.6!}\operatorname{tr} S^4 + \frac{5m^2-20m+14}{8.6!}(\operatorname{tr} S^2)^2 + o(s^4) \quad (2.37)$$

and similarly for $H_2 = \exp(T)$. Also on transformation the ranges of integration become $N(S=0)$ and $N(T=0)$.

It remains to substitute for H_1 and H_2 in $f(H_1,H_2;A,B)$. First we have

$$f(H_1,H_2;A,B) = \operatorname{etr}(AH_1BH_2)(1+F(H_1,H_2;A,B)) \quad (2.38)$$

and

$$\operatorname{tr}(AH_1BH_2) = \sum_{i=1}^{m} a_i b_i h_{ii} k_{ii} + \sum_{i\neq j} a_i b_j h_{ij} k_{ji}.$$

Substituting and simplifying gives

$$\mathrm{tr}(AH_1BH_2) = \sum_{i=1}^{m} a_i b_i - \tfrac{1}{2}\sum_{i<j}\{(a_i b_i + a_j b_j)s_{ij}^2 + (a_i b_i + a_j b_j)t_{ij}^2$$

$$+ 2(a_i b_j + a_j b_i)s_{ij}t_{ij}\} + o(s^2).$$

Each quadratic form can be written in matrix notation as

$$\underline{s}_{ij}' Q_{ij} \underline{s}_{ij} = [s_{ij} \quad b_{ij}] \begin{bmatrix} a_i b_i + a_j b_j & a_i b_j + a_j b_i \\ a_i b_j + a_j b_i & a_i b_i + a_j b_j \end{bmatrix} \begin{bmatrix} s_{ij} \\ t_{ij} \end{bmatrix}$$

so

$$\mathrm{tr}(AH_1BH_2) = \mathrm{tr}(AB) - \tfrac{1}{2}\sum_{i<j} \underline{s}_{ij}' Q_{ij} \underline{s}_{ij} + o(s^2). \qquad (2.39)$$

Substituting in (2.35) using (2.38) and (2.39) gives

$$g(A,B) \simeq \mathrm{etr}(AB) \int_{N(S=0)} \int_{N(T=0)} \exp(-\tfrac{1}{2}\sum_{i<j} \underline{s}_{ij}' Q_{ij} \underline{s}_{ij})(1+o(s^2))$$

$$\times(1+\varphi(S,T;A,B))J(S)J(T)\prod_{i<j} ds_{ij}dt_{ij}. \qquad (2.40)$$

For large values of a_i, b_i the major contribution to the integral is in the region $(S=0)\cap(T=0)$ so we can extend the range of integration for the s_{ij}, t_{ij} to $(-\infty,\infty)$ i.e. integrate over the domains

$$\mathcal{S} = \bigcup_{i<j} \{s_{ij} : -\infty < s_{ij} < \infty\}$$

$$\mathcal{T} = \bigcup_{i<j} \{t_{ij} : -\infty < t_{ij} < \infty\}.$$

Also from (2.37)

$$J(S)J(T) = 1 + \frac{m-2}{24}\mathrm{tr}(S^2+T^2) + o(s^2) \tag{2.41}$$

and substituting in (2.40) we see that the main asymptotic term is given by the evaluation of

$$K(A,B) = \int_{\mathcal{S}}\int_{\mathcal{T}} \exp(-\tfrac{1}{2}\sum_{i<j} \underline{s}'_{ij}\; Q_{ij}\; \underline{s}_{ij}) \prod_{i<j} d\,\underline{s}_{ij}. \tag{2.42}$$

From Appendix 1 section 2 this integral is seen to have the value

$$K(A,B) = \frac{(2\pi)^{\frac{1}{2}m(m-1)}}{\prod_{i<j} c_{ij}^{\frac{1}{2}}} \tag{2.43}$$

where $c_{ij} = (a_i^2-a_j^2)(b_i^2-b_j^2) = (\omega_i-\omega_j)(w_i-w_j)$.

To obtain further terms of the approximation (2.40) it is necessary to evaluate the first few terms of $\varphi(S,T;A,B)$ and $J(S)J(T)$. The expansion to follow is in terms of a_i^{-1} and b_i^{-1} for large a_i, b_i. From (A.1.4) and (2.20) $d_1 = -\frac{1}{8}(n-3)(n-m)$ and it is the coefficient of $C_{(1)}(R^{-1}) = \mathrm{tr}(R^{-1})$ the first term of $F(H_1,H_2;A,B)$, where $R^{-1} = B^{-1}H_1'A^{-1}H_2'$. Setting

$$A^{-1} = \mathrm{diag}(a_i^{-1}) = \mathrm{diag}(\alpha_i) = (\alpha_i\delta_{ij})$$
$$B^{-1} = \mathrm{diag}(b_i^{-1}) = \mathrm{diag}(\beta_i) = (\beta_i\delta_{ij})$$

leads to

$$\mathrm{tr}(B^{-1}H_1'A^{-1}H_2') = \sum_{i,j} \beta_i\alpha_j h_{ji}k_{ij}$$

and on substitution as for (2.39)

$$\mathrm{tr}(B^{-1}H_1'A^{-1}H_2') = \sum_{i=1}^{m} \alpha_i\beta_i - \tfrac{1}{2}\sum_{i<j} \underline{s}_{ij}' P_{ij} \underline{s}_{ij} + o(s^2) \qquad (2.44)$$

where

$$P_{ij} = \begin{bmatrix} \alpha_i\beta_i+\alpha_j\beta_j & \alpha_i\beta_j+\alpha_j\beta_i \\ \alpha_i\beta_j+\alpha_j\beta_i & \alpha_i\beta_i+\alpha_j\beta_j \end{bmatrix}$$

$$= \alpha_i\alpha_j\beta_i\beta_j \; Q_{ij}.$$

Also

$$J(S)J(T) = 1 - \frac{m-2}{12}\sum_{i<j} \underline{s}_{ij}' \, \underline{s}_{ij} + o(s^2). \qquad (2.45)$$

From Appendix 1 section 2 we have

$$g(A,B) \simeq \frac{(2\pi)^{\frac{1}{2}m(m-1)} \mathrm{etr}(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}}} G(A,B) \qquad (2.46)$$

where

$$G(A,B) = 1 - \tfrac{1}{6}(m-2)\sum_{i\neq j} \frac{a_i b_i}{c_{ij}} - \tfrac{1}{8}(n-3)(n-m)\sum_{i=1}^{m} \frac{1}{a_i b_i}$$

$$+ \tfrac{1}{48}(n-3)(n-m)(m-2)\left(\sum_{i\neq j} \frac{a_i b_i + a_j b_j}{a_i b_i \; c_{ij}} + \sum_{i\neq j\neq k} \frac{a_i b_i}{a_k b_k c_{ij}}\right)$$

$$+ \tfrac{1}{16}(n-3)(n-m)\sum_{i<j} \frac{1}{a_i b_i a_j b_j} + \cdots \quad . \qquad (2.47)$$

2.8 Proving the expansion

It is now shown that (2.46) is an asymptotic representation of the integral (2.35). A LEMMA due to HSU [14] is used.

LEMMA 2.3

Let $\varphi(u_1,\ldots u_m)$ and $f(u_1,\ldots u_m)$ be real functions on an m-dimensional closed domain D such that

1. $f > 0$ on D
2. $(f)^n \varphi$ is absolutely integrable over D, $n = 0,1,2,\ldots$
3. all partial derivatives f_{u_i} and $f_{u_i u_j}$ exist and are continuous, $i,j = 1,2,\ldots m$
4. $f(\underline{u})$ has an absolute maximum at an interior point $\underline{\xi}$ of D, so that all $f_{u_i}\big|_{\underline{u}=\underline{\xi}} = 0$ and $\det\left[-f_{u_i u_j}\big|_{\underline{u}=\underline{\xi}}\right] > 0$
5. φ is continuous at $\underline{\xi}$ and $\varphi(\underline{\xi}) \neq 0$.

Then for n large

$$\int_D (f)^n \varphi d\underline{u} \sim \frac{\varphi(\underline{\xi})[f(\underline{\xi})]^n}{[\Delta(\underline{\xi})]^{\frac{1}{2}}} \left[\frac{2\pi}{n}\right]^{\frac{1}{2}m} \qquad (2.48)$$

where $f(\underline{u}) = \exp(\psi(\underline{u}))$ and $\Delta(\underline{u}) = \det[-\psi_{u_i u_j}]$.

This LEMMA is used to prove directly that (2.46) is an asymptotic expansion of the integral.

<u>THEOREM 2.3</u>

Let A,B be $m\times m$ diagonal matrices with

$$a_1 > a_2 > \ldots > a_m > 0$$

$$b_1 > b_2 > \ldots > b_m > 0.$$

Then for A,B large and $g(A,B)$ defined as in (2.35)

$$g(A,B) \sim \frac{(2\pi)^{\frac{1}{2}m(m-1)} \operatorname{etr}(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}}} . \qquad (2.49)$$

<u>Proof</u>

We must show that the conditions of LEMMA 2.3 are satisfied.

Directly after substitution in (2.35) we have

$$g(A,B) = \int_{N(S=0)} \int_{N(T=0)} \operatorname{etr}(AH_1BH_2)(1+\varphi(S,T;A,B))J(S)J(T) \prod_{i<j} ds_{ij}dt_{ij}. \tag{2.50}$$

Now set

$$A = a_1X \qquad x_i = a_1^{-1}a_i$$

$$B = b_1Y \qquad y_i = b_1^{-1}b_i$$

then

$$\operatorname{etr}(AH_1BH_2) = [\operatorname{etr}(XH_1YH_2)]^{a_1b_1}. \tag{2.51}$$

The right hand side is of the form $(f)^n$ where a_1b_1 corresponds to n and $f \equiv \operatorname{etr}(XH_1YH_2)$ is a function of the $m(m-1)$ variables s_{ij}, t_{ij} $i < j$. Also

$$\varphi \equiv (1 + \varphi(S,T;A,B))J(S)J(T)$$

$$\psi \equiv \sum_{i,j} x_iy_jh_{ij}k_{ji} = \operatorname{tr}(XH_1YH_2)$$

and $D \equiv N(S=0) \cap N(T=0)$. Clearly $\underline{\xi}$ corresponds to the point $(S=0,T=0)$ and $(f)^{a_1b_1}$ and hence f have just one maximum in $N(S=0) \cap N(T=0)$. At this point $(S=0,T=0)$

$$(f)^{a_1b_1} = \operatorname{etr}(AB)$$

and from (A.1.9)

$$\varphi(\underline{\xi}) = 1 - \tfrac{1}{8}(n-3)(n-m)\sum_{i=1}^{m} \frac{1}{a_ib_i} + \dots .$$

For large values of A and B we have $\varphi \simeq 1$.

Now by (2.39), with Q_{ij} a function of x_i, y_i

$$\psi = \sum_i x_iy_i - \tfrac{1}{2}\sum_{i<j} \underline{s}'_{ij}Q_{ij}\underline{s}_{ij} + o(s^2).$$

Then

$$\frac{\partial\psi}{\partial \underline{s}_{ij}} = -Q_{ij}\underline{s}_{ij} + o(s). \tag{2.52}$$

Since $$f(S,T) = \exp(\psi(S,T))$$

$$\frac{\partial f}{\partial \underline{s}_{ij}} = \frac{\partial\psi}{\partial \underline{s}_{ij}} \exp(\psi(S,T))$$

and at $(S=0,T=0)$

$$\frac{\partial\psi}{\partial \underline{s}_{ij}} = \frac{\partial f}{\partial \underline{s}_{ij}} = 0.$$

Thus all conditions of LEMMA 2.3 are satisfied and it remains to evaluate $\Delta(S,T)$. Differentiating (2.52) further

$$\frac{\partial^2\psi}{\partial \underline{s}_{ij}\partial \underline{s}_{ij}} = -Q_{ij} + \text{terms of degree at least one}$$

$$\frac{\partial^2\psi}{\partial \underline{s}_{ij}\partial \underline{s}_{uv}} = 0 + \text{terms of degree at least one}$$

$$i < j \qquad u < v$$
$$i \neq u \quad \text{or} \quad j \neq v$$

and

$$\Delta(S,T) = \det\left[-\left.\frac{\partial^2\psi}{\partial \underline{s}_{ij}\partial \underline{s}_{uv}}\right|_{S=0,T=0}\right]$$

$$= \prod_{i<j} \det Q_{ij}$$

$$= \prod_{i<j} (x_i^2-x_j^2)(y_i^2-y_j^2).$$

Hence

$$g(A,B) \sim \frac{\operatorname{etr}(AB).1}{\prod_{i<j}[(x_i^2-x_j^2)(y_i^2-y_j^2)]^{\frac{1}{2}}}\left[\frac{2\pi}{a_1b_1}\right]^{\frac{1}{2}m(m-1)}$$

$$= \frac{(2\pi)^{\frac{1}{2}m(m-1)}\operatorname{etr}(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}}}. \qquad \text{Q.E.D.}$$

Returning to (2.34) we have

$$I \simeq 2^m g(A,B) + 2^m g(A^*,B) \tag{2.53}$$

and it is easy to show that asymptotically the term $g(A^*,B)$ is swamped by $g(A,B)$.

LEMMA 2.4

$$\lim_{a_m, b_m \to \infty} \frac{g(A^*,B)}{g(A,B)} = 0. \tag{2.54}$$

<u>Proof</u>

Substituting the approximations (2.49)

$$\begin{aligned} \frac{g(A^*,B)}{g(A,B)} &\simeq \frac{\operatorname{etr}(A^*B)}{\operatorname{etr}(AB)} \\ &= \exp(-2a_m b_m) \qquad (2.55) \\ &\to 0 \quad \text{as} \quad a_m, b_m \to \infty. \end{aligned}$$

Q.E.D.

Thus from (2.55), (2.53) can be written as

$$I \simeq 2^m g(A,B)(1 + O(\exp(-2a_m b_m))) \quad \text{for large } A,B. \tag{2.56}$$

That is, the effect of $g(A^*,B)$ can be ignored for all practical purposes. To achieve better accuracy it would be better to determine more terms of the series $G(A,B)$ than to include the terms of $g(A^*,B)$. Some more terms are listed in Appendix 1 section 3(A.1.16).

2.9 <u>Summary</u>

From (2.20) and (2.34) remembering that $A^2 = \Omega$, $B^2 = W$

$${}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) \simeq k_6 \det(AB)^{-\frac{1}{2}(n-m)}[g(A,B)+g(A^*,B)] \tag{2.57}$$

and by (2.56) the term $g(A^*,B)$ can be neglected, so we have proved the

THEOREM 2.4

For large values of the non-centrality parameters and of the latent roots w_i, the Bessel function has the asymptotic expansion, for $n \geqslant 2m$, of

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) \sim k \frac{\operatorname{etr}(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}} \det(AB)^{\frac{1}{2}(n-m)}} G(A,B) \tag{2.58} $$

where $a_i^2 = \omega_i$, $b_i^2 = w_i$, $c_{ij} = (a_i^2-a_j^2)(b_i^2-b_j^2)$,

$$ k = \frac{2^{\frac{1}{2}m(n-3)}\Gamma_m(\tfrac{1}{2}m)\Gamma_m(\tfrac{1}{2}n)}{\pi^{\frac{1}{2}m(m+1)}} $$

and

$$ G(A,B) = 1 - \tfrac{1}{6}(m-2)\sum_{i\neq j}\frac{a_i b_i}{c_{ij}} - \tfrac{1}{8}(n-3)(n-m)\sum_{i=1}^{m}\frac{1}{a_i b_i} + o\left(\frac{1}{a^2}\right). \tag{2.59} $$

The numerical assessment of this result is left to Chapter 8, where all the various approximations are examined together. An idea of the range of values of Ω and W for which this is a good approximation will also be specified.

Setting $m=1$ in (2.58) and (2.59) and ignoring all summations involving more than one index gives us the approximation outlined in Chapter 1.

$$ {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}a^2b^2) \sim \frac{2^{\frac{1}{2}(n-3)}\Gamma(\tfrac{1}{2}n)e^{ab}}{\sqrt{\pi}\,(ab)^{\frac{1}{2}(n-1)}}\left(1 - \frac{(n-1)(n-3)}{8ab}\right). \tag{2.60} $$

CHAPTER 3

STATISTICAL APPLICATIONS

3.1 Introduction

In this Chapter we consider an alternative derivation of the asymptotic representation for the Bessel function of two argument matrices using a parameterisation of the Stiefel manifold given by JAMES [23]. The parameterisation is similar to that used for $\mathcal{O}^+(m)$ in Chapter 2, but the substitution can be made without integrating out all implicit variables.

In section 2 the leading (asymptotic) term is derived for the case considered in Chapter 2 i.e. both argument matrices are of full rank, while in section 4 the technique is extended to derive the asymptotic term when one of the matrices is not of full rank.

Two statistical problems are dealt with:

1. In section 3 the non-centrality parameters are estimated by maximum likelihood.
2. In section 5 a likelihood ratio test for the rank of the matrix of means M is derived and its sampling distribution is considered.

The final section ties the results of section 5 to the work of BARTLETT [4], [5] and LAWLEY [26] in deriving multivariate tests of hypothesis.

3.2 The Stiefel manifold

We use as our starting point (2.9) of Chapter 2 i.e.

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) = k_3\int_{O(m)}(dH_1)\int_{V_{nm}}(dH_{\frac{1}{2}})\,\mathrm{etr}\left(\begin{bmatrix} AH_1BH_{\frac{1}{2}} \\ 0 \end{bmatrix}\right) \qquad (3.1)$$

where $k_3 = [\mathrm{Vol}(O(m))\mathrm{Vol}(V_{nm})]^{-1}$.

If $H_{\frac{1}{2}} \in V_{nm}$ is partitioned as $H_{\frac{1}{2}} = [H_{\frac{1}{2}}^{1} \quad H_{\frac{1}{2}}^{2}]$ the integrand of (3.1) becomes $\mathrm{etr}(AH_1BH_{\frac{1}{2}}^{1})$. It is necessary to determine where the maximum of this function occurs.

Since $H_{\frac{1}{2}}^{1}$ is the top left hand corner of an orthogonal matrix, its elements must satisfy the inequalities $(H_2 = (k_{ij})\ \ i,j = 1,\ldots n)$

$$|k_{ij}| \leq 1 \qquad i,j = 1,2,\ldots m.$$

It is easily seen that the maximum of $\mathrm{etr}(AB)$ is attained at the matrices

$$H_1 = H_{\frac{1}{2}}^{1} = \begin{bmatrix} \pm1 & & \\ & \ddots & \\ & & \pm1 \end{bmatrix}$$

giving $H_{\frac{1}{2}}$ the form $[H_{\frac{1}{2}}^{1} \quad 0]$. There are 2^m equal maxima.

It is also clear that we again have the second maximum of $\mathrm{etr}(A^*B)$ but in the following derivations it is ignored.

Since there is an equal contribution from each of the

2^m neighbourhoods of the equal maxima we write (3.1) as

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) \doteq 2^m k_3 \int_{N(I)} (dH_1) \int_{N([I\ 0])} (dH_2^1) \operatorname{etr}(AH_1BH_2^{1\prime}). \qquad (3.2) $$

Again $N(I)$ contains only matrices in $\mathcal{O}^+(m)$ so H_1 can be parameterised. Also JAMES [23] has given a parameterisation of the Stiefel manifold. Thus we can parameterise H_1, H_2^1 by

$$ H_1 = \exp(S) $$

$$ H_2 = \begin{bmatrix} H_2^1 \\ H_2^2 \end{bmatrix} = \exp\left(\begin{bmatrix} T_{11} & T_{12} \\ -T_{12}' & 0 \end{bmatrix}\right) $$

where S, T_{11} are $m\times m$ skew matrices and T_{12} is an $m\times(n-m)$ rectangular matrix.

If we let $H_1 = (h_{ij})$, $S = (s_{ij})$, $T = (t_{ij})$ with $t_{ij} = 0$, i and $j > m$, then writing out the elements

$$ h_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2 + o(s^2) \qquad i = 1,\ldots m $$
$$ h_{ij} = s_{ij} + o(s) \qquad i,j = 1,\ldots m \qquad i \neq j $$
$$ k_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{n} t_{ij}^2 + o(t^2) \qquad i = 1,\ldots m $$
$$ k_{ij} = t_{ij} + o(t) \qquad i,j = 1,\ldots n, \qquad i \neq j. $$

From the integrand of (3.2), neglecting terms of degree greater than 2,

$$ \operatorname{tr}(AH_1BH_2^{1\prime}) = \sum_{i,j=1}^{m} a_i b_j h_{ij} k_{ji} $$
$$ = \sum_{i=1}^{m} a_i b_i (1-\tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2)(1-\tfrac{1}{2}\sum_{j=1}^{n} t_{ij}^2) - \sum_{i,j=1}^{m} a_i b_j s_{ij} t_{ij} $$
$$ = \operatorname{tr}(AB) - \tfrac{1}{2}\sum_{i<j} \underline{s}_{ij}' Q_{ij} \underline{s}_{ij} - \tfrac{1}{2}\sum_{i=1}^{m}\sum_{j=m+1}^{n} a_i b_i t_{ij}^2 \qquad (3.3) $$

where

$$Q_{ij} = \begin{bmatrix} a_ib_i+a_jb_j & a_ib_j+a_jb_i \\ a_ib_j+a_jb_i & a_ib_i+a_jb_j \end{bmatrix} \underline{s}_{ij} = \begin{bmatrix} s_{ij} \\ t_{ij} \end{bmatrix}.$$

For the Jacobian of the transformation

$$(dH_1)(dH_2^1) = (dS)(dT_{11})(dT_{12})(1 + o(s)) \qquad (3.4)$$

where (dS), (dT_{11}), (dT_{12}) stand for $\bigwedge_{i<j}^{m} ds_{ij}$, $\bigwedge_{i<j}^{m} dt_{ij}$ and $\bigwedge_{i=1}^{m} \bigwedge_{j=m+1}^{n} dt_{ij}$ respectively.

Substitute (3.3) and (3.4) in the integrand of (3.2). Since the integrand tends to zero as $|s_{ij}|$, $|t_{ij}|$ tend to ∞, we can change the range of integration to $-\infty < s_{ij} < \infty$, $-\infty < t_{ij} < \infty$ to obtain the leading term of the asymptotic series.

Hence for large values of A and B

$$\int_{N(I)} (dH_1) \int_{N([IO])} (dH_2^1) \operatorname{etr}(AH_1BH_2^{1\prime})$$

$$= \operatorname{etr}(AB) \int_S \int_{T_{11}} \prod_{i<j}^{m} \exp(-\tfrac{1}{2}\underline{s}_{ij}'Q_{ij}\underline{s}_{ij})d\underline{s}_{ij}$$

$$\times \int_{T_{12}} \prod_{i=1}^{m} \prod_{j=m+1}^{n} \exp(-\tfrac{1}{2}a_ib_it_{ij}^2)dt_{ij}$$

$$\simeq \operatorname{etr}(AB) \prod_{i<j} \int_{-\infty}^{\infty}\int \exp(-\tfrac{1}{2}\underline{s}_{ij}'Q_{ij}\underline{s}_{ij})d\underline{s}_{ij}$$

$$\prod_{i=1}^{m} \prod_{j=m+1}^{n} \int_{-\infty}^{\infty} \exp(-\tfrac{1}{2}a_ib_it_{ij}^2)dt_{ij}$$

$$= \frac{\operatorname{etr}(AB)\ (2\pi)^{\frac{1}{2}m(m-1)}\ (2\pi)^{\frac{1}{2}m(n-m)}}{\prod_{i<j} c_{ij}^{\frac{1}{2}}\ \ \det(AB)^{\frac{1}{2}(n-m)}} \tag{3.5}$$

where $c_{ij} = (a_i^2 - a_j^2)(b_i^2 - b_j^2)$.

To summarise we substitute (3.5) in (3.2) to obtain

$${}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) \sim K \frac{\operatorname{etr}(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}} \det(AB)^{\frac{1}{2}(n-m)}} \tag{3.6}$$

where

$$K = \frac{2^{\frac{1}{2}m(n-3)}\Gamma_m(\tfrac{1}{2}m)\Gamma_m(\tfrac{1}{2}n)}{\pi^{\frac{1}{2}m(m+1)}} .$$

This result is the same as the dominant asymptotic term of (2.61). The method of this chapter is much simpler for determining the leading term but it appears to be much more difficult to extend in order to obtain further terms of the series.

3.3 Maximum marginal likelihood estimation

The likelihood factor for the non-central means with known covariance distribution is

$$\operatorname{etr}(-\tfrac{1}{2}\Omega)\,{}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W).$$

We are interested in finding maximum likelihood estimates for $\omega_1,\ldots\omega_m$ from the marginal distribution of $w_1,\ldots w_m$.

Using the asymptotic results, the likelihood function can be factorised as

$$L(\omega_1,\ldots\omega_m) = K(w_1,\ldots w_m)L_1L_2G \tag{3.7}$$

where

$$L_1 = \mathrm{etr}(-\tfrac{1}{2}\Omega)\mathrm{etr}(\Omega W)^{\frac{1}{2}} \tag{3.8}$$

$$L_2 = \det \Omega^{-\frac{1}{4}(n-m)} \prod_{i<j} (\omega_i - \omega_j)^{-\frac{1}{2}} \tag{3.9}$$

G is the asymptotic series (2.47) and K is a function not involving $\omega_1, \ldots \omega_m$.

We begin by finding an estimate $\hat{\omega}_i$ for ω_i using L_1 only and improve it by using L_1L_2. The function G is shown to have negligible effect for large enough values of the w_i. The method of estimation is also due to ANDERSON [1].

Taking (3.8)

$$\ell_1 = \ln L_1 = -\tfrac{1}{2}\sum_{i=1}^{m} \omega_i + \sum_{i=1}^{m} (\omega_i w_i)^{\frac{1}{2}}.$$

Differentiating and equating to zero gives

$$\frac{\partial \ell_1}{\delta \omega_i} = -\tfrac{1}{2} + \tfrac{1}{2}\frac{w_i^{\frac{1}{2}}}{\omega_i^{\frac{1}{2}}} = 0$$

and the approximate maximum likelihood estimate

$$\hat{\omega}_i = w_i. \tag{3.10}$$

Now we consider the effect that the function L_2 has on this estimate. First expand the terms of $\ell_2 = \ln L_2$ in a Taylor series about the points $\hat{\omega}_i$. From (3.9)

$$\ell_2 = -\tfrac{1}{4}(n-m)\sum_{i=1}^{m} \ln \omega_i - \tfrac{1}{2}\sum_{i<j} \ln(\omega_i - \omega_j).$$

Taking the terms separately, with $\omega_i = \hat{\omega}_i + \delta\omega_i$

$$\ln \omega_i = \ln \hat{\omega}_i + \ln\left(1 + \frac{\delta\omega_i}{\hat{\omega}_i}\right)$$

$$= f_i(\hat{\omega}_i) + \frac{\delta\omega_i}{\hat{\omega}_i} + \ldots$$

$$\ln(\omega_i-\omega_j) = g_{ij}(\hat{\omega}_i,\hat{\omega}_j) + \frac{\delta\omega_i - \delta\omega_j}{\hat{\omega}_i - \hat{\omega}_j} + \dots .$$

Combining these results with (3.8) gives

$$\ell = \ln L_1L_2 = -\tfrac{1}{2}\sum_{i=1}^{m}\omega_i + \sum_{i=1}^{m}(\omega_i w_i)^{\frac{1}{2}} - \tfrac{1}{4}(n-m)\sum_{i=1}^{m}\frac{\delta\omega_i}{\hat{\omega}_i}$$

$$-\tfrac{1}{2}\sum_{i<j}\frac{\delta\omega_i - \delta\omega_j}{\hat{\omega}_i - \hat{\omega}_j} + \dots + \Theta(\hat{\omega}_1,\dots\hat{\omega}_m)$$

where Θ is the sum of the functions f_i and g_{ij} and is independent of $\omega_1,\dots\omega_m,\delta\omega_1,\dots\delta\omega_m$.

Differentiating

$$\frac{\partial\ell}{\partial\omega_i} = -\tfrac{1}{2}+\tfrac{1}{2}\frac{w_i^{\frac{1}{2}}}{\omega_i^{\frac{1}{2}}} - \tfrac{1}{4}\frac{n-m}{\hat{\omega}_i} - \tfrac{1}{2}\sum_{j\neq i}\frac{1}{\hat{\omega}_i - \hat{\omega}_j} + \dots .$$

Equating to zero and substituting w_i for $\hat{\omega}_i$,

$$\hat{\omega}_i^{\frac{1}{2}} = w_i^{\frac{1}{2}}\left(1 + \frac{n-m}{2w_i} + \sum_{j\neq i}\frac{1}{w_i-w_j} + \dots\right)^{-1}.$$

Squaring and expanding binomially,

$$\hat{\omega}_i = w_i - (n-m) - 2\sum_{j\neq i}\frac{w_i}{w_i-w_j} + \dots . \qquad (3.11)$$

At this point it would appear that this estimate could be improved by including the effect of G. By considering a numerical example I will show that for large values of ω_i, w_i and for small n the contribution of G is negligible, but not so the correction made by including L_2.

Suppose from a normal sample with $m=3$, $n=10$ sample latent roots of $w_1=100$, $w_2=64$, $w_3=36$ were obtained.

Using (3.11) the estimates for the ω_i are easily calculated.

The exact log likelihood function was approximated by $\ell_1+\ell_2+g$ where $g = \ln G^*$ and G^* is the series (2.47) truncated after four terms. An iterative method was used to locate the function maximum and its coordinates.

Comparing the two sets of results.

	iterative	(3.11)	relative error(%)
ω_1	83.20	84.32	1.35
ω_2	56.24	55.98	0.46
ω_3	33.52	32.70	2.45

All relative errors fall within reasonable bounds.

At these values $\ell = 72.92$, $\ell_1 = 99.47$, $\ell_2 = -26.10$ while $g = -0.45$. Small changes in the ω_i have much greater effect on the values of ℓ_1, ℓ_2 than on g.

Estimation from the marginal distribution is stated by JAMES [22] to lead to unbiassed estimates. He illustrates this by showing that marginal likelihood estimates for the latent roots of the covariance matrix are unbiassed. In order to show the estimates (3.11) are unbiassed estimates of the ω_i it would be necessary to determine $E[w_i]$.

3.4 One argument matrix not of full rank

Using the parameterisation for the Stiefel manifold it is possible to easily derive the leading term of the asymptotic expansion of the Bessel function when one of the argument matrices is not of full rank but all non-zero latent roots are large. In the statistical applications this corresponds to Ω, the matrix of non-centrality parameters, not being of full rank. One application is considered in the next section.

From sections 2.3, 2.4 the integrand of (2.6) can be reduced to $\mathrm{etr}(AH_1BH_2^{1\prime})$ so the equation can be written as

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) = k_1 \int_{O(m)} (dH_1) \int_{O(n)} (dH_2)\ \mathrm{etr}(AH_1BH_2^{1\prime}) \quad (3.12) $$

where $k_1 = [\mathrm{Vol}(O(m))\mathrm{Vol}(O(n))]^{-1}$.

Let the matrix A (and hence $\Omega = A^2$) have rank $k < m$. That is

$$ A = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} k \\ m-k \end{matrix} \qquad A_1 = \mathrm{diag}(a_i). $$

As before we can integrate out over subsets of the orthogonal manifold and since A is of lower rank here, the remaining domains of integration are Stiefel manifolds of lower dimension. Partition the matrices H_1 and $H_2^{1\prime}$ as

$$ H_1 = \underset{m}{\begin{bmatrix} H_1^1 \\ H_1^2 \end{bmatrix}} \begin{matrix} k \\ m-k \end{matrix} \qquad H_2^{1\prime} = \begin{bmatrix} \underset{k}{K^1} & \underset{m-k}{K^2} \end{bmatrix} m . $$

The exponent of the integrand of (3.12) becomes

$$ \mathrm{tr}\left\{\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} H_1^1 \\ H_1^2 \end{bmatrix} B \ [K^1 \quad K^2]\right\} $$

$$= \text{tr}\left\{\begin{bmatrix} A_1H_1^1BK^1 & A_1H_1^1BK^2 \\ 0 & 0 \end{bmatrix}\right\}$$

$$= \text{tr}\,(A_1H_1^1BK^1). \tag{3.13}$$

Now $\text{tr}(A_1H_1^1BK^1)$ contains only elements of $H_1^1 \in \mathcal{V}_{mk}$ and of K where

$$K = \underset{k}{\begin{bmatrix} K^1 \\ K^* \end{bmatrix}}\begin{matrix} m \\ n-m \end{matrix} \in \mathcal{V}_{nk}.$$

We can integrate out over H_1^2 for fixed H_1^1 and over H_2^2 for fixed K by the formulae

$$\int_{H_1^2} (dH_1) = c_1(dH_1^1) \qquad \int_{H_2^2} (dH_2) = c_2(dK)$$

where

$$H_2 = \underset{\quad k \quad\; n-k}{\begin{bmatrix} K & H_2^2 \end{bmatrix}}\, n$$

$$c_1 = \frac{\text{Vol}(\mathcal{O}(m))}{\text{Vol}(\mathcal{V}_{mk})} \qquad c_2 = \frac{\text{Vol}(\mathcal{O}(n))}{\text{Vol}(\mathcal{V}_{nk})}.$$

The equation (3.12) becomes

$${}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) = k_2 \int_{\mathcal{V}_{mk}} (dH_1^1) \int_{\mathcal{V}_{nk}} (dK)\ \text{etr}(A_1H_1^1BK^1) \tag{3.14}$$

where $\quad k_2 = c_1c_2k_1 = [\text{Vol}(\mathcal{V}_{mk})\text{Vol}(\mathcal{V}_{nk})]^{-1}.$

Let $k+q = m$ and partition B into

$$B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}\begin{matrix} k \\ q \end{matrix}$$

where B_1, B_2 are both diagonal matrices. Then $tr(A_1 H_1^1 B K^1)$ has 2^k equal maxima of $tr(A_1 B_1)$ at the matrices

$$H_1^1 = \underset{\quad k \qquad q}{\begin{bmatrix} I^* & 0 \end{bmatrix}} \qquad\qquad K^1 = \underset{k}{\begin{bmatrix} I^* \\ 0 \end{bmatrix}} \begin{matrix} k \\ q \end{matrix}$$

where

$$I^* = \begin{bmatrix} \pm 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \pm 1 \end{bmatrix} .$$

Thus (3.14) becomes approximately

$${}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) \simeq 2^k k_2 \int_{N([I0])} (dH_1^1) \int_{N\left(\begin{bmatrix} I \\ 0 \end{bmatrix}\right)} (dK)\ etr(A_1 H_1^1 B K^1) . \quad (3.15)$$

Now H_1^1 and K can be parameterised by

$$H_1 = \begin{bmatrix} H_1^1 \\ H_1^2 \end{bmatrix} = \exp\left(\begin{bmatrix} S_{11} & S_{12} \\ -S_{12}' & 0 \end{bmatrix}\right)$$

$$H_2 = \begin{bmatrix} K & H_2^2 \end{bmatrix} = \exp\left(\begin{bmatrix} T_{11} & T_{12} \\ -T_{12}' & 0 \end{bmatrix}\right)$$

where S_{11}, T_{11} are $k \times k$ skew matrices and S_{12} $k \times q$ and T_{12} $k \times (n-k)$ are rectangular matrices. With the obvious definitions for S, T and their elements,

$$h_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2 + \cdots \qquad i = 1,\ldots k$$

$$h_{ij} = \quad s_{ij} + \cdots \qquad i = 1,\ldots k,\ j = 1,\ldots m$$

$$k_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{n} t_{ij}^2 + \cdots \qquad i = 1,\ldots k$$

$$k_{ij} = \quad t_{ij} + \cdots \qquad i = 1,\ldots m,\ j = 1,\ldots k.$$

For the integrand of (3.15), neglecting terms of degree greater than 2,

$$\mathrm{tr}(A_1 H_1^{\frac{1}{1}} BK^1) = \sum_{i=1}^{k}\sum_{j=1}^{m} a_i b_j h_{ij} k_{ji}$$

$$= \sum_{i=1}^{k} a_i b_i (1-\tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2)(1-\tfrac{1}{2}\sum_{j=1}^{n} t_{ij}^2) - \sum_{i=1}^{k}\sum_{j=1}^{m} a_i b_j s_{ij} t_{ij}$$

$$= \sum_{i=1}^{k} a_i b_i - \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}\{a_i b_i s_{ij}^2 + a_i b_i t_{ij}^2 + 2a_i b_j s_{ij} t_{ij}\}$$

$$-\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=k+1}^{m}\{a_i b_i s_{ij}^2 + a_i b_i t_{ij}^2 + 2a_i b_j s_{ij} t_{ij}\}$$

$$-\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=m+1}^{n} a_i b_i t_{ij}^2.$$

This expression can be summarised in the notation of quadratic forms as before to facilitate using the standard integral (A.1.5). Let $\underline{s}_{ij}$ and Q_{ij} be as defined in section 2 and

$$Q_{ij}^{*} = \begin{bmatrix} a_i b_i & a_i b_j \\ a_i b_j & a_i b_i \end{bmatrix}.$$

Since $\det Q_{ij}^{*} = a_i^2(b_i^2 - b_j^2) > 0$, Q_{ij}^{*} is positive definite.

Then

$$\operatorname{tr}(A_1 H_1^1 B K^1)$$

$$= \sum_{i=1}^{k} a_i b_i - \tfrac{1}{2}\sum_{i<j}^{k} \underline{s}_{ij}' Q_{ij} \underline{s}_{ij} - \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=k+1}^{m} \underline{s}_{ij}' Q_{ij}^{*} \underline{s}_{ij} - \tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=m+1}^{n} a_i b_i t_{ij}^2 . \qquad (3.16)$$

For the Jacobian of the transformation

$$(dH_1^1)(dK) = (dS_{11})(dS_{12})(dT_{11})(dT_{12})(1 + o(s)) \qquad (3.17)$$

where $(dS_{11}),(dS_{12}),(dT_{11}),(dT_{12})$ stand for $\bigwedge_{i<j}^{k} ds_{ij}$, $\bigwedge_{i=1}^{k}\bigwedge_{j=k+1}^{m} ds_{ij}$, $\bigwedge_{i<j}^{k} dt_{ij}$, $\bigwedge_{i=1}^{k}\bigwedge_{j=k+1}^{n} dt_{ij}$ respectively.

Again the method is to substitute (3.16) and (3.17) in the integral of (3.15) and change the range of integration to $-\infty < s_{ij}, t_{ij} < \infty$ to obtain the asymptotic representation.

$$\int_{N([I0])} (dH_1^1) \int_{N\left(\begin{bmatrix} I \\ 0 \end{bmatrix}\right)} (dK) \operatorname{etr}(A_1 H_1^1 B K^1)$$

$$\simeq \exp\left(\sum_{i=1}^{k} a_i b_i\right) \prod_{i<j}^{k} \int_{-\infty}^{\infty}\!\!\int \exp(-\tfrac{1}{2}\, \underline{s}_{ij}' Q_{ij} \underline{s}_{ij})\, d\underline{s}_{ij}$$

$$\times \prod_{i=1}^{k}\prod_{j=k+1}^{m} \int_{-\infty}^{\infty}\!\!\int \exp(-\tfrac{1}{2}\underline{s}_{ij}' Q_{ij}^{*} \underline{s}_{ij})\, d\underline{s}_{ij}$$

$$\times \prod_{i=1}^{k}\prod_{j=m+1}^{n} \int_{-\infty}^{\infty} \exp(-\tfrac{1}{2} a_i b_i t_{ij}^2)\, dt_{ij}$$

$$= \frac{\exp\left(\sum_{i=1}^{k} a_i b_i\right)(2\pi)^{\frac{1}{2}k(k-1)}(2\pi)^{\frac{1}{2}kq}(2\pi)^{\frac{1}{2}k(n-m)}}{\prod_{i<j}^{k} c_{ij}^{\frac{1}{2}} \prod_{i=1}^{k} \prod_{j=k+1}^{m} [a_i^2(b_i^2-b_j^2)]^{\frac{1}{2}} \prod_{i=1}^{k} (a_i b_i)^{\frac{1}{2}(n-m)}} \ . \tag{3.18}$$

Finally the substitution of (3.18) in (3.15) gives

$${}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) \sim$$

$$\frac{K \ \exp\left(\sum_{i=1}^{k} a_i b_i\right)}{\prod_{i<j} c_{ij}^{\frac{1}{2}} \prod_{i=1}^{k} \prod_{j=k+1}^{m} [a_i^2(b_i^2-b_j^2)]^{\frac{1}{2}} \prod_{i=1}^{k} (a_i b_i)^{\frac{1}{2}(n-m)}} \tag{3.19}$$

where

$$K = \frac{2^{\frac{1}{2}k(n-k-2)}\Gamma_k(\frac{1}{2}m)\Gamma_k(\frac{1}{2}n)}{\pi^{\frac{1}{2}k(m+1)}} \ .$$

For $k=m$ this agrees with (3.6).

3.5 <u>A BARTLETT-LAWLEY type test of rank</u>

The aim is to develop a likelihood ratio test on the rank of the matrix of means M. This is also a test on the number of non-zero non-centrality parameters ω_i. The equivalence of the two follows from

LEMMA 3.1

$$\text{rank }(M) = \text{number of non-zero } \omega_i \ . \tag{3.20}$$

<u>Proof</u>

Since the covariance matrix Σ is positive definite

$$\text{rank }(M) = \text{rank }(MM') = \text{rank }(\Sigma^{-\frac{1}{2}}MM'\Sigma^{-\frac{1}{2}})$$

where $\Sigma^{-\frac{1}{2}}$ is the positive definite square root of Σ^{-1}.

The matrix $\Sigma^{-\frac{1}{2}}MM'\Sigma^{-\frac{1}{2}}$ is symmetric and hence its rank is equal to the number of non-zero latent roots. The matrix $\Sigma^{-1}MM'$ has the same latent roots. Q.E.D.

First we consider the likelihood ratio test of the hypothesis H_0: $M = 0$

against H_1: M arbitrary.

On the alternate hypothesis the likelihood function is

$$L(M) = [2\pi]^{-\frac{1}{2}mn}\det \Sigma^{-\frac{1}{2}n}\operatorname{etr}[-\tfrac{1}{2}\Sigma^{-1}(X-M)(X-M)']$$

and the test statistic is

$$\lambda = \operatorname{etr}(-\tfrac{1}{2}\Sigma^{-1}XX') = \operatorname{etr}(-\tfrac{1}{2}W). \tag{3.21}$$

Now

$$-2 \ln \lambda = \operatorname{tr} W = w_1 +\dots+ w_m \tag{3.22}$$

and putting $X = (x_1 \dots x_n)$ where x_i an $m\times 1$ column vector

$$\operatorname{tr} W = \operatorname{tr} \Sigma^{-1}XX' = \sum_{i=1}^{n} x_i' \Sigma^{-1}x_i .$$

On H_0 each term has a χ^2 distribution on m degrees of freedom so $\operatorname{tr} W$ has a χ^2 distribution on mn degrees of freedom.

A more general hypothesis is now considered. We wish to test H_0 : M has rank $k < m$

against H_1 : M arbitrary.

By LEMMA 3.1 this is equivalent to the test of

$$H_0 : \omega_{k+1} = \dots = \omega_m = 0$$

against H_1 : all $\omega_i > 0$.

We now derive the test statistic and consider its asymptotic distribution. The test statistic used is

$$\Lambda = -2 \ln \lambda = w_{k+1} + \dots + w_m \qquad (3.23)$$

and rather than interrupt the argument at this point, the long but straightforward derivation is given in Appendix 2.

The derivation is a modification of one given by RAO [30]. There he is considering what he calls a test of dimensionality on the matrix of means.

It is interesting that the criterion Λ is derivable, as most BARTLETT-LAWLEY type test statistics for testing these intermediate hypotheses, such as the test that a subset of the latent roots of the covariance matrix are equal, are merely a contraction of the statistic derived for the overall test, such as the sphericity test.

Now by asymptotic theory Λ is distributed as χ^2 where degrees of freedom = number of parameters in H_1

- number of parameters in H_0.

The null hypothesis states that M has rank k. This means that k row vectors of M are linearly independent and the remaining $m-k$ rows are unknown linear combinations of these. Thus H_0 involves $kn + k(m-k)$ parameters. Clearly H_1 involves mn. Hence there are $(m-k)(n-k)$ degrees of freedom for χ^2.

Summarising

THEOREM 3.1

To test H_0 : M has rank $k < m$ against H_1 : M arbitrary, use the statistic

$$\Lambda = w_{k+1} + \dots + w_m .$$

On H_0 $\Lambda \sim \chi^2_d$ where $d = (m-k)(n-k)$.

In the spirit of LAWLEY [26] we can improve our approximation by finding a multiplier c such that $c\Lambda$ is more nearly χ^2_d. A new approach to this problem is given in JAMES [23]. It involves the determination of the conditional distribution of the last q sample roots given the first k.

Now using the asymptotic result (3.19) the joint distribution is given by the

THEOREM 3.2

The asymptotic distribution of the latent roots $w_1, \dots w_k, w_{k+1}, \dots w_m$ depending on the non-centrality parameters $\omega_1, \dots \omega_k$ is

$$f(w_1 \dots w_k, w_{k+1} \dots w_m ; \omega_1, \dots \omega_k) \bigwedge_{i=1}^{m} dw_i$$

$$= \frac{C \exp(-\frac{1}{2}\sum^{k} \omega_i) \exp[\sum^{k} (\omega_i w_i)^{\frac{1}{2}}]}{\prod\limits_{i<j}^{k} [(\omega_i - \omega_j)(w_i - w_j)]^{\frac{1}{2}} \prod\limits_{i=1}^{k} (\omega_i w_i)^{\frac{1}{4}(n-m)}}$$

$$\times \exp(-\tfrac{1}{2}\sum w_i) \prod^{k} w_i^{\frac{1}{2}(n-m-1)} \prod_{i<j}^{k} \left(\frac{w_i - w_j}{\omega_i - \omega_j} \right)^{\frac{1}{2}} \bigwedge_{i=1}^{k} dw_i$$

$$\times \prod_{i=1}^{k} \prod_{j=k+1}^{m} \left[\frac{w_i - w_j}{\omega_i} \right]^{\frac{1}{2}}$$

$$\times \exp\left(-\tfrac{1}{2}\sum_{i=k+1}^{m} w_i\right)\prod_{i=k+1}^{m} w_i^{\frac{1}{2}(n-k-q-1)}\prod_{k<i<j\leq m}(w_i-w_j)\bigwedge_{i=k+1}^{m} dw_i \tag{3.24}$$

where

$$C = \frac{\Gamma_k(\tfrac{1}{2}m)\,\Gamma_k(\tfrac{1}{2}n)}{2^{\frac{1}{2}qn+\frac{1}{2}k(k+2)}\,\pi^{\frac{1}{2}k(m+1)-\frac{1}{2}m^2}\,\Gamma_m(\tfrac{1}{2}n)\,\Gamma_m(\tfrac{1}{2}m)}\ .$$

As in [23] there are two very useful corollaries.

COROLLARY 1

The first k sample roots are asymptotically sufficient for the population roots $\omega_1,\ldots\omega_k$.

COROLLARY 2

The conditional distribution of the last roots $w_{k+1},\ldots w_m$, given the first k, is

$$f_k = f(w_{k+1},\ldots w_m \mid w_1,\ldots w_k;\omega_1,\ldots\omega_k)$$

$$= \text{const.}\ \prod_{i=1}^{k}\ \prod_{j=k+1}^{m}(w_i-w_j)^{\frac{1}{2}}$$

$$\exp\left(-\tfrac{1}{2}\sum_{j=k+1}^{m} w_i\right)\prod_{i=k+1}^{m} w_i^{\frac{1}{2}(n-k-q-1)}\prod_{k<i<j\leq m}(w_i-w_j)\bigwedge_{i=k+1}^{m} dw_i\ . \tag{3.25}$$

and this does not depend on the population parameters $\omega_1,\ldots\omega_k$.

The last line of (3.25) is essentially the null distribution of $w_{k+1},\ldots w_m$ on $n-k$ degrees of freedom. One degree is lost for each variable conditioned on. The test of rank can now be made using this conditional distribution.

If k were zero, then the distribution of the likelihood ratio statistic Λ would be derived from the distribution in the second line of (3.25). The result is

$$\chi^2_d = \Lambda \tag{3.26}$$

where

$$d = qn. \tag{3.27}$$

In testing the last q roots when $k \neq 0$, as a first approximation one could ignore the factor involving $w_1, \ldots w_k$ if these were large. In this case (3.26) would be correct as a first approximation but in (3.27) n is replaced by $(n-k)$ and we have

$$d = q(n-k). \tag{3.28}$$

(Of course (3.27) could be written $q(n-k)$ as $k=0$ in that case.)

By considering the factor involving $w_1, \ldots w_k$ we obtain the refinement of the form $c\Lambda$, which is more nearly χ^2_d. Expanding the product, with $w_k \gg w_{k+1}, \ldots w_m$

$$\prod_{i=1}^{k} \prod_{j=k+1}^{m} (w_i - w_j)^{\frac{1}{2}}$$

$$= \prod_{i=1}^{k}\left(w_i^{\frac{1}{2}q} \prod_{j=k+1}^{m} \left(1 - \frac{w_j}{w_i}\right)^{\frac{1}{2}}\right)$$

$$= \prod_{i=1}^{k}\left(w_i^{\frac{1}{2}q} \prod_{j=k+1}^{m} \left(1 - \frac{1}{2}\frac{w_j}{w_i} + o\left(\frac{1}{w_i}\right)\right)\right)$$

$$= \prod_{i=1}^{k} w_i^{\frac{1}{2}q} \prod_{i=1}^{k}\left(1 - \frac{\Lambda}{2w_i} + o\left(\frac{1}{w_i}\right)\right)$$

$$= \left(1 - \frac{\Lambda}{2}\sum_{i=1}^{k}\frac{1}{w_i} + o\left(\frac{1}{w}\right)\right)\prod_{i=1}^{k} w_i^{\frac{1}{2}q}. \tag{3.29}$$

Using (3.29) the distribution (3.25) is approximated as

$$f_k = \text{const.}\left(1-\tfrac{1}{2}\Lambda\sum^{k}\frac{1}{w_i} + o\left(\frac{1}{w}\right)\right)\prod^{k} w_i^{\frac{1}{2}q}.\text{null distribution} \tag{3.30}$$

where the null distribution is given by

$$\text{const.}\exp\left(-\tfrac{1}{2}\sum_{i=k+1}^{m} w_i\right)\prod_{i=k+1}^{m} w_i^{\frac{1}{2}(n-k-q-1)} \prod_{k<i<j\leq m}(w_i-w_j) \bigwedge_{i=k+1}^{m} dw_i. \tag{3.31}$$

Define the following notation for expectation.

E_0 = expectation with respect to the null distribution (3.31)

E_1 = expectation with respect to the modified distribution (3.30).

To a first approximation, by (3.26) and (3.28)

$$E_0[\Lambda] = d = q(n-k) \tag{3.32a}$$

$$E_0[\Lambda^2] = d(d+2). \tag{3.32b}$$

To find the constant of (3.30) we have

$$1 = E_1[1] = E_0[\text{const.}\left(1-\tfrac{1}{2}\Lambda\sum^{k}\frac{1}{w_i} + o\left(\frac{1}{w}\right)\right)\prod_{i=1}^{k} w_i^{\frac{1}{2}q}]$$

$$= \text{const.}\left(1-\tfrac{1}{2}d\sum^{k}\frac{1}{w_i} + o\left(\frac{1}{w}\right)\right)\prod_{i=1}^{k} w_i^{\frac{1}{2}q}$$

and the modified distribution takes the form

$$f_k = \frac{\left(1 - \tfrac{1}{2}\Lambda\sum\frac{1}{w_i} + o\left(\frac{1}{w}\right)\right)}{\left(1 - \tfrac{1}{2}d\sum\frac{1}{w_i} + o\left(\frac{1}{w}\right)\right)}\ .\ \text{null distribution.} \tag{3.33}$$

The improved multiplier c comes from

$$E_1[\Lambda] = \frac{E_0[\Lambda - \frac{1}{2}\Lambda^2\Sigma \frac{1}{w_i} + o\left(\frac{1}{w}\right)]}{1 - \frac{1}{2}d\Sigma \frac{1}{w_i} + o\left(\frac{1}{w}\right)}$$

$$= [d - \tfrac{1}{2}d(d+2)\Sigma \frac{1}{w_i} + o\left(\frac{1}{w}\right)][1 - \tfrac{1}{2}d\Sigma \frac{1}{w_i} + o\left(\frac{1}{w}\right)]^{-1}$$

$$= d[1 - \overset{k}{\Sigma} \frac{1}{w_i} + o\left(\frac{1}{w}\right)].$$

Thus to order w^{-1} we have

$$\left(1 + \sum_{i=1}^{k} \frac{1}{w_i}\right)\Lambda = \chi^2_d \qquad (3.35)$$

and the results of this section are summarised as

THEOREM 3.3

The statistic

$$\left(1 + \sum_{i=1}^{k} \frac{1}{w_i}\right) \sum_{i=k+1}^{m} w_i \qquad (3.36)$$

is an improved statistic and is asymptotically distributed as χ^2_d.

3.6 Connections with MANOVA and canonical correlations

This problem of rank, or number of non-zero non-centralities, is now shown to be allied to the general MANOVA situation. The results (3.21), (3.22), (3.23), (3.26), (3.27) and (3.28) of the previous section can all be derived as limiting cases.

Let $X_{m\times n}$, $Y_{m\times t}$ be sample matrices on n,t degrees of freedom respectively with the columns all normally and independently distributed with common covariance matrix Σ. Also $E[X] = M$, $E[Y] = 0$.

Thus the likelihood ratio statistic to test $H_0 : M=0$ against $H_1 : M$ arbitrary is

$$\lambda = \left[\frac{\det YY'}{\det(XX'+YY')}\right]^{\frac{1}{2}(n+t)} = \prod_{i=1}^{m}(1-r_i)^{\frac{1}{2}(n+t)} \tag{3.37}$$

where the r_i are the latent roots of

$$\det[XX' - r(XX'+YY')] = 0. \tag{3.38}$$

Asymptotically

$$-2\ln\lambda = -(n+t)\sum_{i=1}^{m}\ln(1-r_i) \sim \chi^2_{mn}. \tag{3.39}$$

The test was proposed by BARTLETT [5] and by considering the expectation of $-2\ln\lambda$ he derived an improved approximation. Allowing for the notational changes $n \to n+t$, $q \to m$, $p \to n$, this has the form

$$-[t - \tfrac{1}{2}(m-n+1)]\sum^{m}\ln(1-r_i) \sim \chi^2_{mn}. \tag{3.40}$$

BARTLETT [4], [5] also proposed that in order to test that M has rank k, the statistic to use is

$$-[t - \tfrac{1}{2}(q-n+1)]\sum_{i=k+1}^{m}\ln(1-r_i) \tag{3.41}$$

which is asymptotically χ^2 on $q(n-k)$ degrees of freedom.

In [26], LAWLEY considers a further adjustment term. Let the f_i be solutions of

$$\det(XX' - fYY') = 0. \tag{3.42}$$

If the f_i estimate population parameters $\beta_i{}^2$ then the approximation is $(1 - r_i = (1 + f_i)^{-1})$

$$[t - \tfrac{1}{2}(q-n+1) + \sum_{i=1}^{k} \frac{1}{f_i}] \sum_{j=k+1}^{m} \ln(1+f_i) \sim \chi_d^2 \tag{3.43}$$

where $d = (m-k)(n-k)$.

Now it was shown in section 1.2 that the non-central means with known covariance distribution can be obtained as a limit from the general non-central means distribution with the substitutions

$$w_i = \frac{r_i t}{1 - r_i} \qquad \text{i.e.} \quad r_i = \frac{w_i}{w_i + t}$$

and then letting $t \to \infty$. Take the right hand side of (3.37), substitute for r_i and let $t \to \infty$ to give

$$\lambda = \prod_{i=1}^{m} (1 + t^{-1}w_i)^{-\frac{1}{2}(n+t)}$$

$$\xrightarrow[t\to\infty]{} \exp(-\tfrac{1}{2}\sum_{i=1}^{m} w_i) = \text{etr}(-\tfrac{1}{2}W)$$

which is the result (3.21).

Similarly substitution for r_i in (3.39) and (3.40) and letting $t \to \infty$ gives (3.22) while the limiting process applied to (3.41) yields (3.23). Under this limiting process the asymptotic distribution of $-2 \log \lambda$ is still a χ^2 on the appropriate number of degrees of freedom, thus yielding (3.26), (3.27) and (3.28).

Also if we substitute $tf_i = w_i$ then as $t \to \infty$

(3.42) becomes $\det(XX' - w\Sigma) = 0$ and (3.43) becomes

$$[1 - \frac{q-n+1}{2t} + \sum_{i=1}^{k} \frac{1}{w_i}] \ln \prod_{i=k+1}^{m} (1 + \frac{w_i}{t})^t$$

$$\xrightarrow[t\to\infty]{} (1 + \sum^{k} \frac{1}{w_i}) \sum_{i=k+1}^{m} w_i$$

which is precisely (3.36).

One final link is with the canonical correlations distribution. As was shown in section 1.2 the general canonical correlations distribution tends to the non-central means with known covariance distribution under the substitutions $w_i = tr_i^2$, $\omega_i = t\rho_i^2$ and taking the limit as $t \to \infty$.

To test the hypothesis $\rho_{k+1} = \cdots = \rho_m = 0$ BARTLETT [4] proposed the statistic $(n \to t,\ p \to m,\ q \to n,\ s \to k)$

$$\chi^2 = -[t - \tfrac{1}{2}(m+n+1)]\log \prod_{i=k+1}^{m} (1-r_i^2) \tag{3.44}$$

where χ^2 has $(m-k)(n-k)$ degrees of freedom. The multiplier was modified by LAWLEY [26] to give an improved approximation. Quoting equation (8) $(n \to t,\ p \to m,\ q \to n)$ this improved multiplier takes the form

$$t - k - \tfrac{1}{2}(m+n+1) + \sum_{i=1}^{k} \frac{1}{r_i^2} . \tag{3.45}$$

Substituting this in (3.44), putting $r_i^2 = t^{-1}w_i$ and letting $t \to \infty$ gives (3.36).

CHAPTER 4

THE BESSEL FUNCTION OF ONE ARGUMENT

4.1 Introduction

Here we apply the reduction process of Chapter 2 to the Bessel function ${}_0F_1(\frac{1}{2}n;\ \frac{1}{4}XX')$ and obtain an asymptotic expansion valid for those X such that XX' has large latent roots. The result is easier to obtain than (2.58) and is made even easier by borrowing freely the results of Chapter 2.

ANDERSON'S integral, in the case $m=k$, is derived as a stage in the process and in section 7 we see how (2.11) can be obtained by averaging (4.4).

Statistically the Bessel function of one argument matrix appears as part of the likelihood factor of the non-central Wishart distribution.

Direct substitution for H in (4.1) again leads to obviously wrong results so some preliminary integrations are in order before setting $H = \exp(S)$.

4.2 ANDERSON'S integral

Directly from JAMES [21] comes the integral

$$ {}_0F_1(\tfrac{1}{2}n;\ \tfrac{1}{4}XX') = k_1 \int_{\mathcal{O}(n)} \operatorname{etr}(XH_1)(dH) \qquad (4.1)$$

and using (2.9), with $H_1 \in \mathcal{V}_{nm}$

$$ {}_0F_1(\tfrac{1}{2}n;\ \tfrac{1}{4}XX') = k_2 \int_{V_{nm}} \mathrm{etr}(XH_1)(dH_1) \qquad (4.2) $$

where $k_1 = [\mathrm{Vol}(O(n))]^{-1}$, $k_2 = [\mathrm{Vol}(V_{nm})]^{-1}$ and $H = [H_1\ H_2] \in O(n)$.

Again this can be shown to be a function of the latent roots of XX' only. Diagonalise X by the transformation of LEMMA 2.1. Set $X = D[A\ 0]E$ with $D \in O(m)$, $E \in O(n)$, $A = \mathrm{diag}(a_i)$ and the $a_i{}^2$ are the latent roots of $\det(XX' - a^2I) = 0$ with $a_1^2 > \ldots > a_m^2 > 0$. Substituting in the integrand of (4.2) gives

$$ \mathrm{etr}(XH_1) = \mathrm{etr}([A\ 0]EH_1D) $$

and it is clear that $EH_1D \in V_{nm}$. Change variables to $K_1 = EH_1D$ and since E,D are constant matrices $(dK_1)=(dH_1)$. Hence (4.2) becomes

$$ {}_0F_1(\tfrac{1}{2}n;\ \tfrac{1}{4}XX') = k_2 \int_{V_{nm}} \mathrm{etr}([A\ 0]K_1)(dK_1). \qquad (4.3) $$

Applying the Stiefel manifold transformation of HERZ,

$$ K_1 = \begin{bmatrix} K_{11} \\ K_{21} \end{bmatrix} = \begin{bmatrix} T \\ U(I_m - T'T)^{\frac{1}{2}} \end{bmatrix} $$

where T is an $m\times m$ real matrix with $T'T \leqslant I$, $U \in V_{n-m,m}$ and there is now the added restriction of $n \geqslant 2m$. For the measures $(dK_1) = \det(I-T'T)^{\frac{1}{2}(n-2m-1)}(dT)(dU)$ and for the integrand of (4.3)

$$\mathrm{etr}\left([A\ 0]\begin{bmatrix}K_{11}\\K_{21}\end{bmatrix}\right) = \mathrm{etr}(AK_{11}) = \mathrm{etr}(AT).$$

So (4.3) becomes, on integrating (dU) over $V_{n-m,m}$

$$ {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}XX') = k_3 \int_{T'T\leqslant I} \mathrm{etr}(AT)\det(I-T'T)^{\frac{1}{2}(n-2m-1)}(dT) \quad (4.4)$$

where

$$k_3 = \frac{\mathrm{Vol}(V_{n-m,m})}{\mathrm{Vol}(V_{nm})} = \frac{\Gamma_m(\frac{1}{2}n)}{\pi^{\frac{1}{2}m^2}\Gamma_m(\frac{1}{2}(n-m))} .$$

Equation (4.4) is ANDERSON'S integral for the case when the matrix X m×n is of rank m. For m=1 this reduces to the POISSON integral (1.30).

The integration over $T'T \leqslant I$ must be reduced to one over $O(m)$. Let $T = HS$ where $H \in O(m)$ and S an m×m symmetric positive definite matrix. From (2.14)

$$(dT) = (dH)\prod_{i<j}(d_i+d_j)(dS) \quad (4.5)$$

where $d_1,\ldots d_m$ are the latent roots of S, and (4.4) becomes

$$ {}_0F_1(\tfrac{1}{2}n;\ \tfrac{1}{4}XX') = k_3 \int_{O(m)} (dH) \int_{S\leqslant I} (dS)\,\mathrm{etr}(AHS)$$

$$\times \det(I-S^2)^{\frac{1}{2}(n-2m-1)} \prod_{i<j}(d_i+d_j). \quad (4.6)$$

The three steps of the classical approximation are now applied. Transform $U = I-S$, U has latent roots

$u_i = 1-d_i$ and (4.6) becomes

$$ {}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}XX') = k_4 \int_{O(m)} (dH)\,\mathrm{etr}(AH) \qquad (4.7a)$$

$$\int_{U\leq I} (dU)\,\mathrm{etr}(-AHU)[\det U(I-\tfrac{1}{2}U)]^{\frac{1}{2}(n-2m-1)} \prod_{i<j}\left(1 - \frac{u_i+u_j}{2}\right) \qquad (4.7b)$$

with $k_4 = 2^{\frac{1}{2}m(n-m-2)}k_3$. Expand binomially to obtain the series (2.17). Apply THEOREM 2.1 term by term to (4.7b) with the series substituted and the range of integration extended to $U > 0$, since the latent roots of A are assumed large. This gives, with $R=AH$, $R^{-1}=H'A^{-1}$

$$ {}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}XX') \simeq k_5 \det A^{-\frac{1}{2}(n-m)}$$

$$\int_{O(m)} \mathrm{etr}(AH)(1 + \sum_{k,\kappa} d_\kappa C_\kappa(R^{-1}))(dH) \qquad (4.8)$$

with d_κ as in (2.20) and $k_5 = \dfrac{2^{\frac{1}{2}m(n-m-2)}\Gamma_m(\frac{1}{2}n)}{\pi^{\frac{1}{2}m^2}}$.

4.3 Finding the maxima

Now (4.8) can be split into integrals over the disjoint subsets $O^+(m)$ and $O^-(m)$. Making use of the device $H^+ = JH$ of Chapter 2, we get

$$\int_{O(m)} f(H;A)(dH) = \int_{O^+(m)} f(H;A)(dH) + \int_{O^+(m)} f(H;A^*)(dH) \qquad (4.9)$$

where f stands for the integrand of (4.8) and $A^* = \mathrm{diag}(a_1 \ldots a_{m-1}, -a_m)$.

To find the stationary points of $tr(AH)$ over $\mathcal{O}^+(m)$ we have, taking differentials and equating to zero,

$$d\ tr(AH) = tr(AH(H'dH)) = 0. \tag{4.10}$$

Using LEMMA 2.2, (4.10) implies that AH is symmetric. Thus $AH = H'A$, or element by element $a_i h_{ij} = h_{ji} a_j$. For $i = 1$,

$$h_{1j} = \frac{a_j}{a_1} h_{j1} \qquad j = 2,\ldots m. \tag{4.11}$$

Since the rows and columns of H are normalised

$$h_{11}^2 + \sum_{j=2}^{m} h_{1j}^2 = h_{11}^2 + \sum_{j=2}^{m} \left(\frac{a_j}{a_1}\right)^2 h_{j1}^2 = 1$$

$$h_{11}^2 + \sum_{j=2}^{m} h_{j1}^2 = 1.$$

By assumption $a_j < a_1$, $j = 2,\ldots m$ and the above are contradictory unless $h_{j1} = 0$, $j = 2,\ldots m$. From (4.11) $h_{1j}=0$, $j = 2,\ldots m$. Thus we can write

$$H = \begin{bmatrix} \pm 1 & 0' \\ 0 & H_1 \end{bmatrix}$$

where $H_1 \in \mathcal{O}(m-1)$, and by repeating the argument on H_1,

$$H = \begin{bmatrix} \pm 1 & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \pm 1 \end{bmatrix}.$$

Hence for $tr(AH)$ the maximum is $tr(A)$ at $H = I$ and for $tr(A^*H)$ the maximum is $tr(A^*)$ at $H = I$.

Unlike the Bessel function of two argument matrices, we have a unique maximum for $tr(AH)$ over $H \in \mathcal{O}^+(m)$ and a unique maximum for $tr(A^*H)$ similarly, rather than the 2^m equal maxima over each subset as before. However we argue as before that for large values of A and A^* the integrands on the right hand side of (4.9) are large only in the neighbourhood of the maximum of $tr(AH)$ and $tr(A^*H)$ respectively. Thus (4.9) approximates to

$$\int_{\mathcal{O}(m)} f(H;A)(dH) \simeq \int_{N(I)} f(H;A)(dH) + \int_{N(I)} f(H;A^*)(dH). \qquad (4.12)$$

4.4 Approximating the integral

We concentrate on

$$g(A) = \int_{N(I)} f(H;A)(dH) \qquad (4.13)$$

but the same procedures may be applied to $g(A^*)$.

Now $N(I)$ contains only $H \in \mathcal{O}^+(m)$ so apply the parameterisation $H = \exp(S)$ where S is an $m\times m$ skew symmetric matrix. Let $H = (h_{ij})$, $S = (s_{ij})$, then

$$h_{ii} = 1 - \tfrac{1}{2}\sum_{j=1}^{m} s_{ij}^2 + o(s^2) \qquad i = 1,\ldots m$$

$$h_{ij} = s_{ij} + o(s) \qquad i \neq j.$$

Also $N(I) \to N(S=0)$ and $(dH) = J(S)\prod_{i<j} ds_{ij}$ where $J(S)$ is given by (2.37).

Substituting for H in $f(H;A) = etr(AH)(1+F(H;A))$.

First

$$\mathrm{tr}(AH) = \sum_{i=1}^{m} a_i \; h_{ii} = \sum_{i=1}^{m} a_i - \tfrac{1}{2}\sum_{i,j} a_i \; s_{ij}^2 + o(s^2)$$

$$= \mathrm{tr}(A) - \tfrac{1}{2}\sum_{i<j} (a_i+a_j)s_{ij}^2 + o(s^2). \qquad (4.14)$$

Thus

$$g(A) \simeq \mathrm{etr}(A) \int_{N(S=0)} \exp(-\tfrac{1}{2}\sum_{i<j} d_{ij} \; s_{ij}^2)(1+\varphi(S;A))J(S) \prod_{i<j} ds_{ij} \qquad (4.15)$$

where $d_{ij} = a_i + a_j$.

For large values of a_i the major contribution to the integral (4.15) comes from integrating over values of s_{ij} near the origin so the range of integration can be extended from $N(S=0)$ to $\bigcup_{i<j} \{s_{ij}: -\infty < s_{ij} < \infty\}$. Furthermore let $A^{-1} = \mathrm{diag}(a_i^{-1}) = \mathrm{diag}(\alpha_i)$, $R^{-1} = (\rho_{ij})$ and

$$F(H;A) = e_1 a_1{}^* + e_2 a_1{}^{*2} + e_3 a_2^* + o(r^{-2}) \qquad (4.16)$$

with the $a_i{}^*$ the elementary symmetric functions of the latent roots of R^{-1}.

Firstly

$$a_1{}^* = \mathrm{tr}(H'A^{-1}) = \sum_{i=1}^{m} \alpha_i h_{ii}$$

$$= \Sigma\, \alpha_i - \tfrac{1}{2}\sum_{i<j} (\alpha_i+\alpha_j)s_{ij}^2 + o(s^2) \qquad (4.17)$$

and secondly $a_1{}^{*2}$ is easily found, while since $\rho_{ij}=\alpha_j h_{ji}$

$$a_2^* = \sum_{i<j}(\rho_{ii}\rho_{jj} - \rho_{ij}\rho_{ji})$$

$$= \sum_{i<j}\alpha_i\alpha_j + \sum_{i<j}\alpha_i\alpha_j s_{ij}^2 - \tfrac{1}{2}\sum_{i<j}\alpha_i\alpha_j\sum_{k=1}^{m}(s_{ik}^2+s_{jk}^2) + o(s^2).$$

Substituting in $F(H;A)$ gives $\varphi(S;A)$ and it is easily seen that all terms derived from (4.15) are evaluable using

$$\int_{-\infty}^{\infty}\exp(-\tfrac{1}{2}ds^2)\,ds = \left[\frac{2\pi}{d}\right]^{\frac{1}{2}}$$

$$\int_{-\infty}^{\infty}s^{2r}\exp(-\tfrac{1}{2}ds^2)\,ds = \left[\frac{2\pi}{d}\right]^{\frac{1}{2}}\frac{1.3\ldots(2r-1)}{d^r} \qquad r = 1,2,\ldots$$

and the integral of an odd power of s gives zero.

Substitution and integration of all terms from (4.15) gives

$$g(A) = \frac{(2\pi)^{\frac{1}{4}m(m-1)}\,\mathrm{etr}(A)}{\prod_{i<j} d_{ij}^{\frac{1}{2}}}\,G(A) \tag{4.18}$$

and stopping at terms of second degree in $\frac{1}{a_i}$

$$G(A) = 1 - \tfrac{1}{8}(n-3)(n-m)\sum\frac{1}{a_i} - \tfrac{1}{12}(m-2)\sum_{i<j}\frac{1}{a_i+a_j}$$

$$+ \tfrac{1}{128}(n-3)(n-5)(n-m)(n-m+2)\sum\frac{1}{a_i^2}$$

$$+ \tfrac{1}{128}(n-m)(n^3-n^2m+2n^2+8m^2-10nm+23n-5m-2)\sum_{i<j}\frac{1}{a_ia_j}$$

$$+ \tfrac{1}{128}(n-3)(n-m)(m-2)\sum_{i<j}\sum_{k=1}^{m}\frac{1}{a_k(a_i+a_j)}$$

$$+ \frac{(m-2)(5m-11)}{2.6!}\sum_{i<j}\frac{3}{(a_i+a_j)^2} + \frac{5m^2-23m+38}{6!}\sum_{\substack{i<j\\ j<k}}\sum_{i<k}\frac{1}{(a_i+a_j)(a_i+a_k)}$$

$$+ \frac{5m^2-20m+14}{8.6!}\sum_{\substack{i<j\\ i<k}}\sum_{k<\ell}\frac{1}{(a_i+a_j)(a_k+a_\ell)} + o\left(\frac{1}{a^2}\right). \tag{4.19}$$

4.5 Proving the approximation

Using HSU'S LEMMA we can show that (4.18) is an asymptotic expansion for the integral (4.13) for large values of A.

THEOREM 4.1

Let $A = \text{diag}(a_i)$ be $m \times m$ with $a_1 > a_2 > \ldots > a_m > 0$. Then for A large and g(A) defined as in (4.13)

$$g(A) \sim \frac{(2\pi)^{\frac{1}{4}m(m-1)} \operatorname{etr}(A)}{\prod_{i<j} (a_i + a_j)^{\frac{1}{2}}} . \qquad (4.20)$$

Proof

The proof follows the lines of that for THEOREM 2.3. Set $A = a_1 X$, $x_i = a_1^{-1} a_i$, $\operatorname{etr}(AH) = [\operatorname{etr}(XH)]^{a_1}$ etc...Q.E.D.

Again it is easy to show that

$$g(A^*) = O(\exp(-2a_m) g(A))$$

indicating the relative unimportance of the second term.

4.6 Summary

From (4.8) and (4.13)

$${}_0F_1(\tfrac{1}{2}n; \tfrac{1}{4}XX') \simeq k_5 \det A^{-\frac{1}{2}(n-m)} [g(A) + g(A^*)] . \qquad (4.21)$$

The results may be summarised in the

THEOREM 4.2

Let the matrix XX' have latent roots a_i^2. Then for large values of the a_i^2 the Bessel function has the asymptotic representation

$$ {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}XX') \simeq \frac{k \ \mathrm{etr}(A)}{\prod_{i<j}(a_i+a_j)^{\frac{1}{2}} \det A^{\frac{1}{2}(n-m)}} G(A) \qquad (4.22) $$

where $G(A)$ is given by (4.19) and $k = \dfrac{2^{\frac{1}{4}m(2n-m-5)}\Gamma_m(\frac{1}{2}n)}{\pi^{\frac{1}{4}m(m+1)}}$.

Setting $m=1$ this reduces to

$$ {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}x^2) \simeq \frac{2^{\frac{1}{2}(n-3)}e^x}{\sqrt{\pi}\, x^{\frac{1}{2}(n-1)}}\left(1 - \frac{(n-1)(n-3)}{8x} + \frac{(n+1)(n-1)(n-3)(n-5)}{128x^2} \cdots\right) $$

which agrees with the first few terms of (1.31).

Again numerical evaluations are left for Chapter 8.

4.7 The averaged ANDERSON'S Integral.

The integral (2.11) can be deduced directly from (4.4) by averaging over the orthogonal group. Take the left hand side of (2.11)

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}A^2,B^2) = c_1\int_{O(m)} {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}A^2HB^2H')(dH) \qquad (4.23) $$

where $c_1 = [\mathrm{Vol}(O(m))]^{-1}$. The argument A^2HB^2H' has the same latent roots as $BH'A(BH'A)'$ so (4.23) becomes

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}A^2,B^2) = c_1\int_{O(m)} {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}BH'A(BH'A)')(dH) $$

$$ \underset{(4.4)}{=} c_2\int_{O(m)}\int_{R'R\leq I} \mathrm{etr}(BH'AR)\det(I-R'R)^{\frac{1}{2}(n-2m-1)}(dR)(dH) $$

where $c_2 = \dfrac{\Gamma_m(\frac{1}{2}m)\Gamma_m(\frac{1}{2}n)}{2^m\pi^{m^2}\Gamma_m(\frac{1}{2}(n-m))}$. Making the substitution $T'=R$ and using the fact that $\mathrm{tr}(XY) = \mathrm{tr}(X'Y')$ with $X = BH'A$, $Y=T'$ gives (2.11).

CHAPTER 5

THE P.D.E. ASYMPTOTIC FORMULA FOR ${}_0F_1$

5.1 Introduction

The result of Chapter 4 for the Bessel function of one matrix argument was given in terms of inverse powers of the latent roots and was asymptotic on these becoming large. In this Chapter we consider an asymptotic expansion for ${}_0F_1(\frac{1}{2}n;R)$ (R $m\times m$ symmetric) where the series is given in powers of n^{-1} on the condition that the matrix R depends on n.

The asymptotic expansion is derived using a system of partial differential equations given by MUIRHEAD [27]. The system is a generalisation of that given by JAMES [17] for ${}_0F_1(\frac{1}{2}m;R)$ and many of the results used in section 2 are taken from that paper.

Finally the expansion is related to the ${}_0F_1$ appearing in the likelihood factor of the non-central Wishart distribution.

5.2 Using the differential equations

Let R be an $m\times m$ complex symmetric matrix with latent roots $R_1, R_2, \ldots R_m$. Then from MUIRHEAD [27]

THEOREM 5.1

The function ${}_0F_1(\frac{1}{2}n;R)$ is the unique solution of each of the m differential equations

$$R_i\frac{\partial^2 F}{\partial R_i^2} + \left\{\tfrac{1}{2}n-\tfrac{1}{2}(m-1) + \tfrac{1}{2}\sum_{j\neq i} \frac{R_i}{R_i-R_j}\right\}\frac{\partial F}{\partial R_i} - \tfrac{1}{2}\sum_{j\neq i} \frac{R_j}{R_i-R_j}\frac{\partial F}{\partial R_j} = F$$

$$i = 1,2,\ldots m \qquad (5.1)$$

subject to the conditions that

(a) F is symmetric in $R_1,R_2,\ldots R_m$, and

(b) F is analytic about $R=0$ and $F(0)=1$.

In statistical applications the matrix R is restricted to being positive semi-definite. That is all $R_i \geq 0$, $i = 1,2,\ldots m$. However for the expansion an even more restrictive condition is needed. Let R have the form

$$R = nS \quad \text{for each } n, \qquad (5.2)$$

where S is a fixed $m\times m$ symmetric matrix.

Thus we can determine the behaviour of ${}_0F_1$ as $n\to\infty$.

LEMMA 5.1

$$\lim_{n\to\infty} {}_0F_1(\tfrac{1}{2}n;R) = \operatorname{etr}(2S) \qquad (5.3)$$

Proof

Expand ${}_0F_1$ in a zonal series and take limits. Since the series is absolutely convergent, the order of the operations of summation and taking limits can be reversed. Thus, substituting $R=nS$

$$\lim_{n\to\infty} {}_0F_1(\tfrac{1}{2}n;R) = \sum_{k,\kappa} \lim_{n\to\infty} \frac{C_\kappa(nS)}{(\tfrac{1}{2}n)_\kappa k!} .$$

Now $C_\kappa(nS) = n^k C_\kappa(S)$ and taking limits the individual terms reduce to $C_\kappa(2S)/k!$. Summation of these terms gives the required result. Q.E.D.

First we obtain a system of PDE's in the latent roots $S_1, S_2, \ldots S_m$ of S. Make the transformation (5.2). Then $R_i = nS_i$ and differentiating $\frac{\partial F}{\partial R_i} = \frac{1}{n}\frac{\partial F}{\partial S_i}, \frac{\partial^2 F}{\partial R_i^2} = \frac{1}{n^2}\frac{\partial^2 F}{\partial S_i^2}$

The system (5.1) becomes

$$S_i\frac{\partial^2 F}{\partial S_i^2} + \left\{\tfrac{1}{2}n - \tfrac{1}{2}(m-1) + \tfrac{1}{2}\sum_{j\neq i}\frac{S_i}{S_i - S_j}\right\}\frac{\partial F}{\partial S_i} - \tfrac{1}{2}\sum_{j\neq i}\frac{S_j}{S_i-S_j}\frac{\partial F}{\partial S_j} = nF$$

$$i = 1,2,\ldots m. \qquad (5.4)$$

Now using (5.3), for large values of n the function ${}_0F_1$ can be factorised in the form

$$F = \operatorname{etr}(2S)G \qquad (5.5)$$

and we get PDE's for G. Differentiating partially in (5.5)

$$\frac{\partial F}{\partial S_i} = \operatorname{etr}(2S)\frac{\partial G}{\partial S_i} + 2\operatorname{etr}(2S)G$$

$$\frac{\partial^2 F}{\partial S_i^2} = \operatorname{etr}(2S)\frac{\partial^2 G}{\partial S_i^2} + 4\operatorname{etr}(2S)\frac{\partial G}{\partial S_i} + 4\operatorname{etr}(2S)G.$$

Substituting in (5.4) and cancelling an etr(2S) gives the system

$$S_i\frac{\partial^2 G}{\partial S_i^2} + \left\{\tfrac{1}{2}n - \tfrac{1}{2}(m-1) + 4S_i + \tfrac{1}{2}\sum_{j\neq i}\frac{S_i}{S_i-S_j}\right\}\frac{\partial G}{\partial S_i} - \tfrac{1}{2}\sum_{j\neq i}\frac{S_j}{S_i-S_j}\frac{\partial G}{\partial S_j} + 4S_iG = 0$$

$$i = 1,2,\ldots m. \qquad (5.6)$$

By the condition (a) on the solution of (5.1), F is a symmetric function. Also etr(2S) is a symmetric function. Hence from (5.5) G is also a symmetric function. This suggests that we try a series expansion in elementary symmetric functions. That is, a solution of the form

$$G = 1 + \sum_{u=1}^{\infty} \frac{P_u(a)}{n^u} \tag{5.7}$$

where the $P_u(a)$ are polynomials in the elementary symmetric functions $a_1, a_2, \ldots a_m$ of the variables $S_1, S_2, \ldots S_m$.

JAMES [17] has shown how to transform from a PDE in the variables to a PDE in the elementary symmetric functions of them. Let $a_j^{(i)}$ for $j = 1,2,\ldots m-1$ denote the jth elementary symmetric function of the variables $S_1 \ldots S_m$ omitting S_i. Introducing the dummy variables

$a_0 = a_0^{(i)} = 1$

$a_j = 0 \qquad j = -1,-2,\ldots \text{and } m+1, m+2, \ldots$

$a_j^{(i)} = 0 \qquad j = -1,-2,\ldots \text{and } m, m+1, \ldots.$

we have the relationship

$$a_j = S_i\, a_{j-1}^{(i)} + a_j^{(i)} \qquad -\infty < j < \infty \qquad i = 1,2,\ldots m. \tag{5.8}$$

The partial derivative formulae are

$$\frac{\partial}{\partial S_i} = \sum_{\nu=1}^{m} a_{\nu-1}^{(i)} \frac{\partial}{\partial a_\nu}$$

$$\frac{\partial^2}{\partial S_i^2} = \sum_{\nu,\mu=1}^{m} a_{\nu-1}^{(i)}\, a_{\mu-1}^{(i)} \frac{\partial^2}{\partial a_\nu \partial a_\mu}.$$

On substituting for the partial derivatives and applying (5.8), (5.6) gives

$$\sum_{1\leq\mu\leq\nu\leq m} (2-\delta_{\nu\mu})\Big(\sum_{j=1}^{\mu} a_{\mu+\nu-1} a_{j-1}^{(i)} - \sum_{j=1}^{\mu} a_{\mu-j}\, a_{\nu+j-1}^{(i)}\Big)\frac{\partial^2 G}{\partial a_\nu \partial a_\mu}$$

$$+ \tfrac{1}{2}\sum_{j=1}^{m} (n+1-j) a_{j-1}^{(i)} \frac{\partial G}{\partial a_j} + 4\sum_{j=1}^{m} (a_j - a_j^{(i)})\frac{\partial G}{\partial a_j} = 4(a_1^{(i)} - a_1)G$$

$$i = 1,2,\ldots m. \tag{5.9}$$

To proceed we need the following

LEMMA 5.2 ([17] p 371)

If $S_1, S_2, \ldots S_m$ are indeterminates and $a_1, a_2 \ldots a_m$ the elementary symmetric functions of them, and $a_1^{(1)} \ldots a_{m-1}^{(1)}$ the elementary symmetric functions of $S_1 \ldots S_m$ with S_1 omitted and if $\lambda_0(a), \lambda_1(a), \ldots \lambda_{m-1}(a)$ are functions of $a_1 \ldots a_m$ such that

$$\lambda_0(a) + \lambda_1(a)a_1^{(1)} + \ldots + \lambda_{m-1}(a)a_{m-1}^{(1)} = 0$$

then $\lambda_0(a) = 0,\ \lambda_1(a) = 0 \ldots \lambda_{m-1}(a) = 0.$

Using the LEMMA 5.2 each of the m PDE's in (5.9) can be expanded into a system of m PDE's. Each member of (5.9) will give the same system of derived PDE's. Equating the coefficients of $a_{j-1}^{(1)}$ to zero for $j = 1, 2, \ldots m$ we have

$$\sum_{\nu,\mu=1}^{m} c_{\mu\nu}^{(j)}(S_1 \ldots S_m)\frac{\partial^2 G}{\partial a_\mu \partial a_\nu} + \tfrac{1}{2}(n+1-j)\frac{\partial G}{\partial a_j}$$

$$+ 4\sum_{\nu=1}^{m} a_\nu \frac{\partial G}{\partial a_\nu}\delta_{1j} + 4a_1 G\delta_{1j} = 4\frac{\partial G}{\partial a_{j-1}}(1-\delta_{1j}) + 4G\delta_{2j}$$

$$j = 1, 2, \ldots m \qquad (5.10)$$

where $c_{\mu\nu}^{(j)} = c_{\nu\mu}^{(j)}$ and for $\mu \leqslant \nu$

$$c_{\mu\nu}^{(j)} = \begin{cases} a_{\mu+\nu-j} & 1 \leqslant j \leqslant \mu \\ 0 & \mu < j \leqslant \nu \\ -a_{\mu+\nu-j} & \nu < j \leqslant \mu+\nu \\ 0 & \mu+\nu < j \end{cases} \qquad j = 1, 2, \ldots m.$$

Written more explicitly, the system of differential equations (5.10) is

$j = 1$: $(a_0^{(1)})$

$$\operatorname{tr}\left\{\begin{bmatrix} a_1 & a_2 & \cdots & a_m \\ a_2 & & \cdot & \\ \vdots & \cdot & & \\ a_m & & & 0 \end{bmatrix}\begin{bmatrix} \frac{\partial^2 G}{\partial a_1^2} & \cdots & \frac{\partial^2 G}{\partial a_1 \partial a_m} \\ \vdots & \ddots & \\ \frac{\partial^2 G}{\partial a_m \partial a_1} & & \frac{\partial^2 G}{\partial a_m^2} \end{bmatrix}\right\} \qquad (5.11a)$$

$$+ \tfrac{1}{2}n \frac{\partial G}{\partial a_1} + 4 \sum_{\nu=1}^{m} a_\nu \frac{\partial G}{\partial a_\nu} + 4a_1 G = 0$$

$j = 2$: $(a_1^{(1)})$

$$\operatorname{tr}\left\{\begin{bmatrix} -1 & 0 & \cdots & & 0 \\ 0 & a_2 & a_3 & \cdots & a_m \\ \cdot & a_3 & & \cdot & \\ \vdots & \vdots & \cdot & & \\ 0 & a_m & & & 0 \end{bmatrix}\left(\frac{\partial^2 G}{\partial a_\nu \partial a_\mu}\right)\right\} \qquad (5.11b)$$

$$+ \tfrac{1}{2}(n-1) \frac{\partial G}{\partial a_2} - 4 \frac{\partial G}{\partial a_1} = 4G.$$

For the general PDE $(a_{j-1}^{(1)})$

$$\operatorname{tr}\left\{\begin{bmatrix} 0 & & & 0 & -1 & & & \\ & & & -1 & -a_1 & & & \\ & & \cdot & & & & 0 & \\ & \cdot & & & & & & \\ 0 & -1 & & & -a_{j-3} & & & \\ -1 & -a_1 & & -a_{j-3} & -a_{j-2} & & & \\ & & & & & a_j & a_{j+1} & \cdots\, a_m \\ & 0 & & & & a_{j+1} & & \cdot \\ & & & & & \vdots & \cdot & \\ & & & & & a_m & & 0 \end{bmatrix}\left(\frac{\partial^2 G}{\partial a_\nu \partial a_\mu}\right)\right\}$$

(5.11c)

$$+ \tfrac{1}{2}(n+1-j)\frac{\partial G}{\partial a_j} - 4\frac{\partial G}{\partial a_{j-1}} = 0$$

$j = 3 \ldots m.$

Substitute for G in (5.11) using (5.7) and equate coefficients of powers of n^{-1}.

Term independent of $n(u=0)$.

$$j = 1: \quad \tfrac{1}{2}\frac{\partial P_1(a)}{\partial a_1} + 4a_1 = 0 \tag{5.12a}$$

$$j = 2: \quad \tfrac{1}{2}\frac{\partial P_1}{\partial a_2} = 4 \tag{5.12b}$$

$$j > 2: \quad \tfrac{1}{2}\frac{\partial P_1}{\partial a_j} = 0. \tag{5.12c}$$

Hence $P_1(a)$ is a function of a_1 and a_2 alone and must have the form

$$P_1(a) = c_0 + c_1a_1 + c_2a_2 + c_3a_1{}^2. \tag{5.13}$$

Condition (b) on (5.1) is needed here to provide a unique solution to the equations (5.12). If S=0, then ${}_0F_1(\tfrac{1}{2}n;nS) = 1$ and hence G=1. Thus $P_1(0)=0$ and $c_0=0$. Substituting (5.13) in the equations (5.12a) and (5.12b) gives the solution

$$P_1(a) = 8a_2 - 4a_1{}^2. \tag{5.14}$$

Next the coefficient of n^{-1}.

$$j = 1: \quad \tfrac{1}{2}\frac{\partial P_2}{\partial a_1} = 8a_1 + 32a_1{}^2 - 32a_2 - 32a_1a_2 + 16a_1{}^3$$

$$j = 2: \quad \tfrac{1}{2}\frac{\partial P_2}{\partial a_2} = -4 - 32a_1 + 32a_2 - 16a_1{}^2$$

$$j = 3: \quad \tfrac{1}{2}\frac{\partial P_2}{\partial a_3} = 32 \qquad j > 3: \quad \frac{\partial P_2}{\partial a_j} = 0.$$

Solving the system under condition (b)

$$P_2(a) = 32a_2^2-32a_2a_1^2+8a_1^4+64a_3+64a_1a_2 + \tfrac{64}{3}a_1^3-8a_2+8a_1^2. \tag{5.15}$$

By considering successively the coefficients of $n^{-2},n^{-3},\ldots.$ the polynomials $P_3(a),P_4(a),\ldots$ can be found. The number of terms per polynomial increases sharply with u as $P_u(a)$ is a polynomial of degree 2u. Rather than evaluate further $P_u(a)$, we now consider a series expansion in another set of symmetric functions that gives fewer terms per polynomial.

5.3 The direct method

Let us consider a series of the form

$$G = 1 + \sum_{u=1}^{\infty} \frac{Q_u(r)}{n^u} \tag{5.16}$$

where the $Q_u(r)$ are polynomials in the power sums $r_1, r_2 \ldots$ of the variables $S_1 \ldots S_m$. This is analogous to (5.7).

Substituting (5.16) directly into (5.6) gives

$$S_i \Sigma \frac{1}{n^u} \frac{\partial^2 Q_u}{\partial S_i^2} + \left\{\tfrac{1}{2}n - \tfrac{1}{2}(m-1) + 4S_i + \tfrac{1}{2}\sum_{j\neq i} \frac{S_i}{S_i - S_j}\right\} \Sigma \frac{1}{n^u} \frac{\partial Q_u}{\partial S_i}$$

$$-\tfrac{1}{2}\sum_{j\neq i} \frac{S_j}{S_i - S_j} \Sigma \frac{1}{n^u} \frac{\partial Q_u}{\partial S_j} + 4S_i\left\{1 + \Sigma \frac{Q_u}{n^u}\right\} = 0$$

$$i = 1, 2, \ldots m. \tag{5.17}$$

Term independent of $n (u=0)$

$$\tfrac{1}{2} \frac{\partial Q_1}{\partial S_i} + 4S_i = 0 \qquad i = 1, 2 \ldots m.$$

Solving

$$Q_1(r) = -4S_i^2 + \text{similar terms} \quad i = 1, \ldots m.$$

Combining all m results and expressing the solution in terms of power sums

$$Q_1(r) = -4r_2. \tag{5.18}$$

The coefficient of $\frac{1}{n}$,

$$\frac{\partial Q_2}{\partial S_i} = 16S_i + 8\sum_{j\neq i} S_j + 32S_i r_2 + 64S_i^2 \qquad i = 1, \ldots m.$$

Integrating

$$Q_2(r) = 8S_i^2 + 8\sum_{j\neq i} S_iS_j + 8S_i^4 + 16S_i^2\sum_{j\neq i} S_j^2 + \tfrac{64}{3}S_i^3 \quad (5.19)$$

$$+ \text{ similar terms.}$$

Combining all m results

$$Q_2(r) = 8r_2 + 8\sum_{i<j} S_iS_j + 8r_4 + 16\sum_{i<j} S_i^2S_j^2 + \tfrac{64}{3}r_3. \quad (5.20)$$

It is easily seen that

$$r_1^2 = r_2 + 2\sum_{i<j} S_iS_j$$

$$r_2^2 = r_4 + 2\sum_{i<j} S_i^2S_j^2$$

and substitution in (5.20) gives the answer in power sums as

$$Q_2(r) = 8r_2^2 + \tfrac{64}{3}r_3 + 4r_2 + 4r_1^2. \quad (5.21)$$

For general u, the m equations have the form

$$\frac{\partial Q_{u+1}}{\partial S_i} = -2S_i\frac{\partial^2 Q_u}{\partial S_i^2} + (m-1)\frac{\partial Q_u}{\partial S_i} - 8S_i\frac{\partial Q_u}{\partial S_i} - 8S_iQ_u$$

$$-\sum_{j\neq i}\frac{S_i\frac{\partial Q_u}{\partial S_i} - S_j\frac{\partial Q_u}{\partial S_j}}{S_i - S_j}$$

$$i = 1,\ldots m. \quad (5.22)$$

5.4 The two methods, a comparison

From JAMES [21] comes a table of zonal polynomials in terms of power sums and elementary symmetric functions. By equating the two

$$a_1 = r_1$$

$$2a_2 = r_1^2 - r_2$$

$$6a_3 = r_1^3 - 3r_1r_2 + 2r_3.$$

Substitution in P_1 and P_2 shows that

$$P_u(a) = P_u(a(r)) = Q_u(r) \quad u = 1,2 \qquad (5.23)$$

but Q_1 and Q_2 only involve half as many terms as P_1 and P_2. From this it is inferred that in general Q_u will contain less terms than P_u but one can be converted into the other if necessary.

The method of Section 3 is more direct as no change of variables is required in the PDE(5.6) before substituting (5.16). On the negative side the main disadvantages appear to be the evaluation of the term

$$\sum_{j \neq i} \frac{S_i \dfrac{\partial Q_u}{\partial S_i} - S_j \dfrac{\partial Q_u}{\partial S_j}}{S_i - S_j} \qquad (5.24)$$

and the combination of m results like (5.19) into the polynomial (5.20) where the method is one of trial and error. The solution of (5.22) introduces functions that are not power sums or even elementary symmetric functions e.g. $\sum_{i<j} S_i{}^2 S_j{}^2$.

For the first method the main problem is to write down the system of equations for each P_u. The general equation (5.10) is not as simple as (5.22) because the coefficients $c_{\mu\nu}^{(j)}$ depend on the equation, but the solving is trivial since we need only work in terms of elementary symmetric functions.

The main asset of the first method is that it would be quite simple to write a computer programme to evaluate the polynomials $P_u(a)$ recursively. For the second method, the term (5.24) seems to need human intervention to handle. Also there is this problem of mixed terms and their conversion to power sums.

5.5 Statistical applications

The non-central Wishart distribution involves the Bessel function ${}_0F_1(\frac{1}{2}n;\frac{1}{4}\Sigma^{-1}MM'\Sigma^{-1}XX')$ where n is the number of degrees of freedom of the sample matrix $X m\times n$, $E[X] = M$ and the columns of X are normally and independently distributed with common covariance matrix Σ.

Let $XX' = nS$, then $E[S] = \Sigma + \frac{1}{n}MM'$ and S is clearly bounded in probability as $n \to \infty$. For n large S can be treated as a constant matrix. Thus XX' can be considered as a function of n alone for n sufficiently large, satisfying the condition of LEMMA 5.1.

The Bessel function can be written as ${}_0F_1(\frac{1}{2}n;nT)$ where $T = \frac{1}{4}\Sigma^{-1}MM'\Sigma^{-1}S$. Approximating for n large gives

$$ {}_0F_1(\tfrac{1}{2}n;nT) \simeq \operatorname{etr}(2T)G(n;T) \tag{5.25} $$

where G can be expanded as in (5.7) or (5.16).

This approximation would be particularly applicable to power function calculations. There we are dealing with small deviations from the central distribution. In particular, it should be most useful for situations involving only one non-zero latent root or perhaps more generally a small number of them non-zero.

The approximation could also be used to evaluate the likelihood function of the non-central Wishart distribution, viz.,

$$L(M,\Sigma) \propto \operatorname{etr}(-\tfrac{1}{2}\Sigma^{-1}MM')\,{}_0F_1(\tfrac{1}{2}n;nT). \qquad (5.26)$$

Its usefulness would be limited to that part of the range for which the latent roots of nT are small or the sample size would need to be large.

No numerical calculations were done to determine the region of application of the approximation. This would be an extensive study in itself.

CHAPTER 6

THE LAGUERRE POLYNOMIAL $L_k^a(S)$

6.1 Introduction

All previous asymptotic expansions are only valid for large n or for argument matrices with large latent roots. For small values the zonal series converges rapidly enough. The problem is that for "medium" values the asymptotic expansions do not work and the convergence of the zonal series is too slow.

The aim of this Chapter is to work with the zonal series for the $_0F_1$ functions and by rearrangement of series obtain more rapid convergence. The series will be rearranged in terms of the generalised Laguerre polynomials introduced by HERZ [12] and CONSTANTINE [8]. Two such rearrangements will be demonstrated. Both are applicable to the one and two argument Bessel.

A similar Laguerre type expansion exists for the $_1F_1$. This is involved in the non-central moments of the generalised variance and the likelihood ratio statistic. The expansion is included for reasons of completeness only.

The numerical work is left to Chapter 8. All matrices referred to in the following sections are $m \times m$.

6.2 The classical results

Let us first review the classical formulae for functions of a single variable. Referring to Chapter 10, Section 12 of ERDELYI ET AL [9] we see that the Laguerre polynomials $\mathcal{L}_n^a(x)$ are defined as

$$\mathcal{L}_n^a(x) = \sum_{m=0}^{n} \binom{n+a}{n-m} \frac{(-x)^m}{m!} \qquad n = 0,1,2,\ldots \quad a > -1 \tag{6.1}$$

The identities to be generalised are

$$\sum_{n=0}^{\infty} \frac{\mathcal{L}_n^a(x)}{\Gamma(n+a+1)} z^n = (xz)^{-\frac{1}{2}a} e^z J_a[2(xz)^{\frac{1}{2}}] \tag{6.2}$$

and

$$\sum_{n=0}^{\infty} \frac{n!\mathcal{L}_n^a(x)\mathcal{L}_n^a(y)}{(a+1)_n} z^n = (1-z)^{-a-1} \exp\left[-\frac{z(x+y)}{1-z}\right] {}_0F_1(a+1;\frac{xyz}{(1-z)^2}). \tag{6.3}$$

Both can be proved by applying the Laplace transform

$$g(w) = \int_0^{\infty} e^{-vw} v^a f(v) dv. \tag{6.4}$$

Apply to z in (6.2) and y in (6.3) and in both cases they reduce to the main generating function for Laguerre polynomials i.e.

$$\sum_{n=0}^{\infty} \mathcal{L}_n^a(x) z^n = (1-z)^{-a-1} \exp\left[\frac{xz}{z-1}\right] \qquad |z| < 1. \tag{6.5}$$

As is well known (and easily verified) the functions J and ${}_0F_1$ are related by

$$(\tfrac{1}{2}z)^{-a} J_a(z) = \frac{1}{\Gamma(a+1)} {}_0F_1(a+1;-\tfrac{1}{4}z^2). \tag{6.6}$$

A different normalisation is used in [8] so that for $m=1$ the generalised Laguerre polynomial reduces to a multiple of (6.1) i.e.

$$L_n^a(x) = n!\, \mathcal{L}_n^a(x). \tag{6.7}$$

Using (6.6) and (6.7) the identities (6.2), (6.3) and (6.5) can be written in a form suitable for generalisation $(\Gamma(n+a+1) = \Gamma(a+1)(a+1)_n)$

$$\sum_{n=0}^{\infty} \frac{L_n^a(x) z^n}{(a+1)_n n!} = e^z \, {}_0F_1(a+1;-xz) \tag{6.8}$$

$$\sum_{n=0}^{\infty} \frac{L_n^a(x) L_n^a(y)}{(a+1)_n n!} z^n = (1-z)^{-a-1} \exp\left[-\frac{z(x+y)}{1-z}\right] {}_0F_1\left(a+1;\frac{xyz}{(1-z)^2}\right)$$

$$|z| < 1 \tag{6.9}$$

$$\sum_{n=0}^{\infty} \frac{L_n^a(x)}{n!} z^n = (1-z)^{-a-1} \exp\left[\frac{xz}{z-1}\right] \qquad |z| < 1. \tag{6.10}$$

6.3 <u>The matrix generalisations</u>

The following definitions and THEOREM 6.1 are taken from [8]. Let S be a positive definite symmetric matrix, $p = \frac{1}{2}(m+1)$, $a > -1$ and κ a partition of k, then the generalised Laguerre polynomial $L_\kappa^a(S)$ has the definition (corresponding to (6.1))

$$L_\kappa^a(S) = (a+p)_\kappa C_\kappa(I) \sum_{n=0}^{k} \sum_{\nu} \frac{(-1)^n \binom{\kappa}{\nu}}{(a+p)_\nu} \frac{C_\nu(S)}{C_\nu(I)}. \tag{6.11}$$

The "binomial" coefficients $\binom{\kappa}{\nu}$ are defined by

$$\frac{C_\kappa(I+S)}{C_\kappa(I)} = \sum_{n=0}^{k} \sum_{\nu} \binom{\kappa}{\nu} \frac{C_\nu(S)}{C_\nu(I)}. \tag{6.12}$$

In Chapter 7 specific methods for calculating the $\binom{\kappa}{\nu}$ to any order will be considered.

Now the analogue of (6.10).

THEOREM 6.1 ([8] Theorem 1)

The generating function for the Laguerre polynomials is

$$\det(I-Z)^{-a-p}\int_{O(m)} \operatorname{etr}(-SH'Z(I-Z)^{-1}H)(dH) = \sum_{k,\kappa} \frac{L_\kappa^a(S)}{k!}\frac{C_\kappa(Z)}{C_\kappa(I)}$$

$(x \to S,\ z \to Z)$ $\qquad ||Z|| < 1 \qquad (6.13)$

or alternatively

$$\sum_{k,\kappa} \frac{L_\kappa^a(S)\ C_\kappa(Z)}{k!\ C_\kappa(I)} = \det(I-Z)^{-a-p}\ {}_0F_0^{(m)}(S, Z(Z-I)^{-1})$$

$||Z|| < 1 \qquad (6.14)$

where Z is a complex symmetric matrix and $||Z||$ denotes the maximum of the absolute values of the latent roots of Z. To see that (6.14) is a matrix generalisation of (6.10) write

$$\exp\left[\frac{xz}{z-1}\right] = {}_0F_0(xz(z-1)^{-1}) \ . \qquad (6.15)$$

The two theorems that follow are natural generalisations of (6.8) and (6.9). Both are proved by use of the Laplace transform and their Laplace transforms are shown to reduce to (6.14). As in the single variable case, if two functions have equal Laplace transforms then the functions are equal. We use (cf.(6.4))

$$g(W) = \int_{Z>0} \operatorname{etr}(-ZW)\det Z^a f(Z)(dZ) \qquad (6.16)$$

where W is a complex symmetric matrix. The appropriate theory is covered in [7].

In the proofs two standard Laplace transforms are used:

1. CONSTANTINE [7]

$$\int_{Z>0} \operatorname{etr}(-ZW)\det Z^{b-p}C_\kappa(Z)(dZ) = \Gamma_m(b,\kappa)\det W^{-b}C_\kappa(W^{-1}) \tag{6.17}$$

2. CONSTANTINE [8]

$$\int_{Z>0} \operatorname{etr}(-ZW)\det Z^{a}L_\kappa^a(Z)(dZ) = \Gamma_m(a+p,\kappa)\det W^{-a-p}C_\kappa(I-W^{-1}). \tag{6.18}$$

The generalisation of (6.8).

THEOREM 6.2

For $S > 0$, $Z > 0$, $a > -1$

$$\sum_{k,\kappa} \frac{L_\kappa^a(S) \quad C_\kappa(Z)}{(a+p)_\kappa k!C_\kappa(I)} = \operatorname{etr}(Z)\,{}_0F_1^{(m)}(a+p;S,-Z) \tag{6.19}$$

$(x \to S,\ z \to Z)$

<u>Proof</u>

Apply (6.16) to both sides. The left hand side becomes

$$\sum_{k,\kappa} \frac{L_\kappa^a(S)}{(a+p)_\kappa k!C_\kappa(I)} \int_{Z>0} \operatorname{etr}(-ZW)\det Z^a C_\kappa(Z)(dZ)$$

$$\underset{(6.17)}{=} \sum_{k,\kappa} \frac{L_\kappa^a(S)}{(a+p)_\kappa k!C_\kappa(I)}\Gamma_m(a+p,\kappa)\det W^{-a-p}C_\kappa(W^{-1})$$

$$\underset{(6.14)}{=} \Gamma_m(a+p)[\det W(I-W^{-1})]^{-a-p}\,{}_0F_0^{(m)}(S,W^{-1}(W^{-1}-I)^{-1}).$$

Expanding ${}_0F_1^{(m)}$ in its zonal series and applying (6.16) the right hand side becomes

$$\sum_{k,\kappa} \frac{(-1)^k C_\kappa(S)}{(a+p)_\kappa k! C_\kappa(I)} \int_{Z>0} \operatorname{etr}(-Z(W-I)) \det Z^a C_\kappa(Z)(dZ)$$

$$\underset{(6.17)}{=} \sum_{k,\kappa} \frac{(-1)^k C_\kappa(S)}{(a+p)_\kappa k! C_\kappa(I)} \Gamma_m(a+p,\kappa) \det(W-I)^{-a-p} C_\kappa((W-I)^{-1})$$

$$= \quad \Gamma_m(a+p) \det(W-I)^{-a-p} {}_0F_0^{(m)}(S,(I-W)^{-1}).$$

Since $I-W = (W^{-1}-I)W$, both sides are equal. Q.E.D.

Now the generalisation of (6.9).

THEOREM 6.3

For $S > 0$, $Z > 0$, $a > -1$

$$\sum_{k,\kappa} \frac{L_\kappa^a(S) L_\kappa^a(Z)}{(a+p)_\kappa k! C_\kappa(I)} t^k$$

$$= (1-t)^{-m(a+p)} \operatorname{etr}\left(-\frac{t}{1-t}(S+Z)\right) {}_0F_1^{(m)}\left(a+p; \frac{t}{(1-t)^2}S, Z\right) \quad (6.20)$$

$(x \to S,\ y \to Z,\ z \to t)$ $\qquad |t| < 1.$

<u>Proof</u>

Apply (6.16) to both sides. The left hand side becomes

$$\sum_{k,\kappa} \frac{L_\kappa^a(S)\ t^k}{(a+p)_\kappa k! C_\kappa(I)} \int_{Z>0} \operatorname{etr}(-ZW) \det Z^a L_\kappa^a(Z)(dZ)$$

$$\underset{(6.18)}{=} \sum_{k,\kappa} \frac{L_\kappa^a(S)\ t^k}{(a+p)_\kappa k! C_\kappa(I)} \Gamma_m(a+p,\kappa) \det W^{-a-p} C_\kappa(I-W^{-1})$$

$$= \Gamma_m(a+p) \det W^{-a-p} \sum_{k,\kappa} \frac{L_\kappa^a(S) C_\kappa(t(I-W^{-1}))}{k! C_\kappa(I)}$$

$$\underset{(6.13)}{=} \Gamma_m(a+p) [\det W(I-t(I-W^{-1}))]^{-a-p} \int_{O(m)} \operatorname{etr}[SH' t(I-W^{-1})(t(I-W^{-1})-I)^{-1} H](dH)$$

The right hand side becomes

$$(1-t)^{-m(a+p)}\operatorname{etr}\left(-\frac{t}{1-t}S\right)\sum_{k,\kappa}\frac{C_\kappa\left(\frac{t}{(1-t)^2}S\right)}{(a+p)_\kappa k!C_\kappa(I)}$$

$$\int_{Z>0}\operatorname{etr}(-Z(W+\frac{t}{1-t}I))\det Z^a C_\kappa(Z)(dZ)$$

$$\underset{(6.17)}{=}(1-t)^{-m(a+p)}\operatorname{etr}(-\frac{t}{1-t}S)\sum_{k,\kappa}\frac{C_\kappa\left(\frac{t}{(1-t)^2}S\right)}{(a+p)_\kappa\ k!C_\kappa(I)}$$

$$\Gamma_m(a+p,\kappa)\det(W+\frac{t}{1-t}I)^{-a-p}C_\kappa((W+\frac{t}{1-t}I)^{-1})$$

$$=\Gamma_m(a+p)\det[(1-t)W+tI]^{-a-p}$$

$$\int_{O(m)}\operatorname{etr}(-\frac{t}{1-t}S)\operatorname{etr}\left[\frac{t}{(1-t)^2}SH'(W+\frac{t}{1-t}I)^{-1}H\right](dH).$$

The exponent in the integral reduces to

$$\frac{t}{1-t}SH'[-I+((1-t)W+tI)^{-1}]H.$$

Comparing the two sides for the terms in det we have $W(I-t(I-W^{-1})) = (1-t)W+tI$. Finally we must show that

$$t(I-W^{-1})(t(I-W^{-1})-I)^{-1} = \frac{t}{1-t}[((1-t)W+tI)^{-1}-I].$$

Taking the right hand side gives

$$\frac{t}{1-t}(W-t(W-I))^{-1}[I-(1-t)W-tI]$$

$$= t(I-t(I-W^{-1}))^{-1}(W^{-1}-I). \quad \text{Q.E.D.}$$

6.4 The non-central moments of the likelihood ratio statistic

The non-central moments of the generalised variance $\det(XX')$ were given in [12] as

$$E[\det(XX')^k] = 2^{km}\frac{\Gamma_m(k+\frac{1}{2}s)}{\Gamma_m(\frac{1}{2}s)}\det\Sigma^k{}_1F_1^{(m)}(-k;\tfrac{1}{2}s;-\tfrac{1}{2}\Omega) \qquad (6.21)$$

where XX' $m\times m$ has the non-central Wishart on s degrees of freedom and noncentrality $\Omega = \text{diag}(\omega_i)$ and the ω_i the latent roots from $\det(MM' - \omega\Sigma) = 0$. If XX' is as above and YY' $m\times m$ is a central Wishart on t degrees of freedom then the non-central moments of the likelihood ratio statistic have been given in [7] as

$$E\left[\left(\frac{\det YY'}{\det(XX'+YY')}\right)^k\right] =$$

$$= \frac{\Gamma_m(k+\frac{1}{2}t)\Gamma_m(\frac{1}{2}(s+t))}{\Gamma_m(\frac{1}{2}t)\Gamma_m(k+\frac{1}{2}(s+t))}\,{}_1F_1^{(m)}(k;k+\tfrac{1}{2}(s+t);-\tfrac{1}{2}\Omega). \qquad (6.22)$$

A classical formula for the ${}_1F_1$ in Laguerre polynomials is given by RAINVILLE [29] as

$$(1-t)^{-c}{}_1F_1(c;1+a;\frac{xt}{t-1}) = \sum_{n=0}^{\infty}\frac{(c)_n L_n^a(x)t^n}{(1+a)^n\ n!}\ . \qquad (6.23)$$

This generalises to a formula given by JAMES [21] equation (138).

THEOREM 6.4

$$\det(I-Z)^{-c}{}_1F_1^{(m)}(c;a+p;S,Z(Z-I)^{-1})$$
$$= \sum_{k,\kappa}\frac{(c)_\kappa L_\kappa^a(S)C_\kappa(Z)}{(a+p)_\kappa k!\ C_\kappa(I)} \qquad (6.24)$$

$(t \to Z,\ x \to S)$.

<u>Proof</u>

The proof follows by applying the Laplace transform to S. Rearrangements required are the same as those of

[8] Theorem 1. Q.E.D.

Using the trick (1.17) gives $S=I$, $Z(Z-I)^{-1} = -\frac{1}{2}\Omega$, and $Z = \Omega(\Omega + 2I)^{-1}$. No attempt is made to evaluate (6.21) and (6.22) numerically.

CHAPTER 7

CALCULATION OF ZONAL POLYNOMIALS AND $\binom{\kappa}{\nu}$

7.1 Introduction

Recently JAMES [24] developed an expansion of zonal polynomials in terms of the monomial symmetric functions $M_\kappa(S)$ (defined by (7.1)). Sections 2 and 3 deal with the numerical calculation of the Bessel functions using the zonal series. In section 3 an algorithm for the recursive evaluation of the $M_\kappa(S)$ is stated and proved.

Also, as stated in Chapter 6, in order to proceed with the evaluation of the $L_\kappa^a(S)$ a method for calculating the $\binom{\kappa}{\nu}$ to any order is needed. In section 4 a formula for them, in terms of the product coefficients $g_{\nu\mu}^\kappa$ (defined by (7.21)), is presented. Of course it then follows that we need to determine the $g_{\nu\mu}^\kappa$ and a method using the monomial symmetric function expansion is given in the following section.

Finally section 6 contains some summation identities for the $g_{\nu\mu}^\kappa$ and $\binom{\kappa}{\nu}$ which could prove fruitful if studied further.

7.2 The $Z_\kappa(\cdot)$ in terms of the $M_\kappa(\cdot)$

The fundamental units of the theory are the zonal polynomials $Z_\kappa(S)$. As yet there is no known direct formula for them. JAMES [24] found a partial differential equation satisfied by the $Z_\kappa(S)$ and showed how to use this

P.D.E to find expansions of the zonal polynomials in terms of the monomial symmetric functions (msf's) $M_\kappa(S)$.

First the definition of $M_\kappa(S)$. As usual S is an $m\times m$ symmetric matrix with latent roots $s_1, s_2 \ldots s_m$. Let $\kappa = (k_1 k_2 \ldots k_r) = (1^{\pi_1} 2^{\pi_2} \ldots i^{\pi_i} \ldots)$ $r \leqslant m$. Then

$$M_\kappa(S) = \sum_{i,j\ldots u} s_i^{k_1} s_j^{k_2} \ldots s_u^{k_r} \quad i,j\ldots u = 1,2,\ldots m \tag{7.1}$$

where the summation is over all distinct $i,j\ldots u$ for which each distinct term appears once only. The number of terms in the sum is

$$\frac{[m]_r}{\pi_1!\,\pi_2!\ldots\pi_i!\ldots} \tag{7.2}$$

where $[m]_r = m(m-1)\ldots(m-r+1)$ $([m]_r = (-1)^r(-m)_r)$.

The P.D.E. for $Z_\kappa(S)$ is

$$\sum_{i=1}^{m} s_i^2 \frac{\partial^2}{\partial s_i^2} Z_\kappa(S) + \sum_{i\neq j}^{m} \frac{s_i^2}{s_i - s_j} \frac{\partial}{\partial s_i} Z_\kappa(S) - \sum_{i=1}^{m} k_i(k_i+m-i-1) Z_\kappa(S) = 0 \tag{7.3}$$

where κ is a partition with at most m nonzero parts.

My task here was to write a computer programme using the recurrence relations derived from (7.3) to find the coefficients $c_{\kappa\tau}$ of the expansion

$$Z_\kappa(S) = \sum_{\tau\leqslant\kappa} c_{\kappa\tau} M_\tau(S) \tag{7.4}$$

where κ, τ are partitions of k and the ordering is defined in section 1.4. I found it feasible to compute all values of $c_{\kappa\tau}$ for $k \leqslant 13$.

The formula for the leading coefficient is $(k_{r+1}=0)$

$$c_{\kappa\kappa} = 2^k \prod_{\ell=1}^{r} \prod_{i=1}^{\ell} (\tfrac{1}{2}\ell - \tfrac{1}{2}(i-1) + k_i - k_\ell)_{k_\ell - k_{\ell+1}} \tag{7.5}$$

and the recurrence relation is

$$c_{\kappa\tau} = \sum_{\tau<\mu\leqslant\kappa} \frac{[(\ell_i+t)-(\ell_j-t)]c_{\kappa\mu}}{\rho_\kappa - \rho_\tau} \tag{7.6}$$

where

$$\rho_\tau = \sum_{i=1}^{s} \ell_i(\ell_i-1) \qquad \tau = (\ell_1 \ldots \ell_s) \tag{7.7}$$

and $\mu = (\ell_1 \ldots \ell_i + t \ldots \ell_j - t \ldots \ell_s)$ for $t = 1,2,\ldots \ell_j$ such that when the elements of μ are arranged in descending order the inequality $\tau < \mu \leqslant \kappa$ is satisfied.

Tables for $k = 6,7$ (Table 7.1) and $k = 8$ (Table 7.2) are given. A table for $k = 9$ would have 31 rows and columns with some of the entries having 9 digits. This would most certainly require at least two pages and seems to me to be too vast to cope with by hand anyway.

The tables in [24] are for $k = 1,..5$ and do not go far enough to reveal a very interesting point. For $k = 6$ there are two zero entries in the table. These zeros are easily explained. For example

$$Z_{(41^2)}(S) = 270\, M_{(41^2)}(S) + 0\, M_{(3^2)}(S) + \ldots.$$

Note that in the lexicographic ordering a partition of 2 parts follows one of 3 parts. If S has 2 non-zero latent roots then clearly $M_{(41^2)}(S) = 0$ and by theory $Z_{(41^2)}(S) = 0$ but $M_{(3^2)}(S) \neq 0$ hence its coefficient must κ = 6

Z_κ \ M_κ	6	51	42	411	33	321	3111	222	2211	21111	111111	$\chi_{[2\kappa]}(1)$
6	10395	5670	4725	3150	4500	2700	1800	2430	1620	1080	720	1
51		1050	600	1170	540	852	1008	648	828	816	720	54
42			432	288	288	432	504	648	576	648	720	275
411				270	0	162	468	108	348	576	720	616
33					720	432	288	288	528	576	720	132
321						112	168	168	288	456	720	2673
3111							144	0	96	360	720	1925
222								360	240	360	720	462
2211									120	288	720	2640
21111										168	720	1485
111111											720	132

κ = 7

Z_κ \ M_κ	7	61	52	511	43	421	4111	331	322	3211	31111	2221	22111	211111	1111111	$\chi_{[2\kappa]}(1)$
7	135135	72765	59535	39690	55125	33075	22050	31500	28350	18900	12600	17010	11340	7560	5040	1
61		11340	6300	12390	5400	8700	10350	8100	6120	7980	7920	6480	6660	6000	5040	77
52			3600	2400	2160	3408	4032	2592	4392	3984	4680	4536	4608	4920	5040	637
511				2310	0	1320	3870	1188	792	2652	4464	1944	3492	4560	5040	1365
43					2880	1728	1152	3312	2592	3024	2880	3456	3888	4320	5040	1001
421						648	972	432	792	1364	2280	1776	2568	3720	5040	12012
4111							900	[illegible]	0	540	2088	360	1512	3240	5040	7644
331								1008	672	1232	1344	1344	2352	3360	5040	6435
322									840	560	840	1550	1920	3000	5040	9009
3211										320	768	480	1272	2640	5040	42042
31111											648	0	432	2040	5040	14014
2221												720	1080	2160	5040	12012
22111													504	1680	5040	21450
211111														960	5040	7007
1111111															5040	429

TABLE 7.1 The $c_{\kappa\tau}$

M_κ

$\kappa = 8$ (Z_κ)	8	71	62	611	53	521	5111	44	431	422	4211	41111	332	3311	3221	32111	311111	2222	22211	221111	2111111	11111111	$\chi_{[2\kappa]}(1)$
8	2 27025	1081080	873180	582120	793800	476280	317520	771750	441000	396900	264600	176400	378000	252000	226800	151200	100800	204120	136080	40720	60480	40320	1
71		145530	79380	156870	66150	107730	124520	63000	96750	72900	95850	95400	67500	90000	72900	75000	68400	58320	63180	58320	49680	40320	104
62			37800	25200	21600	34800	41400	19440	24480	42936	39024	40296	35424	32976	38976	40752	44040	44064	40608	41424	41760	40320	1260
611				24570	0	13650	40320	0	11700	7800	26550	45000	7020	24480	17880	32760	43200	12951	25380	34800	39600	40320	2640
53					21600	12960	8640	14400	20448	15552	19872	19584	26352	23904	27216	27048	29520	25921	31530	33696	36720	40320	3640
521						5280	7920	0	3168	6072	10512	17784	4752	8064	11616	17208	25920	12960	18576	25296	33120	40320	38220
5111							7560	0	0	0	4320	16920	0	3888	2592	11232	24480	1728	8064	18576	30240	40320	23100
44								40320	23040	20736	13824	9216	13824	20496	20736	24192	23040	31104	27648	31104	34560	40320	1430
431									3888	2592	4752	5184	4752	9024	8416	12768	16320	9600	15216	21696	29520	40320	68640
422										4536	3024	4536	3024	2016	6496	8232	13440	13344	13488	18384	27360	40320	60060
4211											1800	432	0	1200	2200	5456	12720	2400	7224	14208	25200	40320	262080
41111												3960	0	0	0	2376	11520	0	1584	8208	21600	40320	76440
332													5040	3360	6160	6720	8400	6720	11760	10800	25200	40320	51480
3311														2688	1792	4992	7680	2688	6144	13056	23040	40320	150150
3221															1600	2400	4000	3840	6360	11040	20880	40320	336336
32111																1296	4320	0	1944	7008	18000	40320	480480
311111																	3600	0	0	2400	13680	40320	91728
2222																		6640	5760	8640	17280	40320	24024
22211																			2520	6048	15120	40320	171600
221111																				2688	11520	40320	150150
2111111																					6480	40320	32032
11111111																						40320	1430

TABLE 7.2 The $c_{\kappa\tau}$

be zero.

By a generalisation of the above it is easy to prove

LEMMA 7.1

If partition κ of r parts is followed in the lexicographic list by a partition τ of less than r parts, then $c_{\kappa\tau} = 0$.

7.3 Zonal series and evaluating the $M_\kappa(S)$

The Bessel functions have the zonal series definitions

$$ {}_0F_1(c;R) = \sum_{k,\kappa} \frac{C_\kappa(R)}{(c)_\kappa k!} \tag{7.8}$$

$$ {}_0F_1^{(m)}(c;R,S) = \sum_{k,\kappa} \frac{C_\kappa(R)C_\kappa(S)}{(c)_\kappa C_\kappa(I)k!} \tag{7.9}$$

where $$C_\kappa(S) = c(\kappa) Z_\kappa(S)$$

and $$c(\kappa) = \frac{\chi_{[2\kappa]}(1)}{1.3\ldots(2k-1)}$$. The character $\chi_{[2\kappa]}(1)$ is easily calculated using a modified form of the expression given by JAMES [22] for $C_\kappa(I)$. Its consideration is left to Appendix 3. From (7.4) it is clear that with $c_{\kappa\tau}$ known everything depends on the evaluation of the $M_\kappa(S)$.

Direct evaluation of (7.1) by summing over all $[m]_r$ permutations and dividing by the repetition factor $\pi_1!..\pi_1!...$ is a very tedious process. What is needed is an algorithm for building up a table of the $M_\kappa(S)$ by expressing each msf in terms of msf's of lower degree.

The clue is given by VAN DER WAERDEN [32] in exercises 5,6 on page 82.

Quoting directly with $(m) = \sum_{i=1}^{h} x_i^m$

"5. Let

$$(k_1 \ldots k_h) = \Sigma\, x_1^{k_1} x_2^{k_2} \ldots x_h^{k_h}$$

with the summation performed on all distinct permuted terms which may be obtained if we take the order of the subscripts different from 1,2,...h. Prove that

$$(k_1 \ldots k_h)(m) = c_1(k_1+m, k_2 \ldots k_h) + c_2(k_1, k_2+m, \ldots k_h) + \ldots + c_h(k_1, k_2 \ldots k_h+m) + c_0(k_1 \ldots k_h, m) \qquad (7.10)$$

where the coefficients $c_i (i=1,\ldots h)$ and c_0 indicate how many of the integers in the symbols to which they belong are equal to k_i+m and to m, respectively.

6. Solve the formula found in Ex.5 for $(k_1 \ldots k_h, m)$..."

The equation (7.10) is not strictly correct but a corrected version is given by (7.11). Let $\kappa = (k_1 \ldots k_r) = (1^{\pi_1} \ldots i^{\pi_i} \ldots)$ and also let $(k_1 \ldots k_r)$ indicate the msf of m variables associated with the partition κ. Abbreviate

r part msf ≡ msf associated with r part partition.

The one part msf or power sum r_t can be written $M_{(t)}$ or just (t).

Considering the product of M_κ and r_t, (7.10) is replaced by

THEOREM 7.1

$$(k_1\ldots k_r)(t) = c_1(k_1+t,k_2\ldots k_r)+c_2(k_1,k_2+t,\ldots k_r)+\ldots\ldots$$
$$+\ c_q(k_1\ldots k_q+t\ldots k_r)+\ldots$$
$$+\ c_r(k_1\ldots k_r+t)+c_0(k_1\ldots k_r,t) \qquad (7.11)$$

where

$$c_q = \frac{\pi_i+1}{\pi_j} \qquad \begin{array}{l} k_q = j \\ k_q+t = i \end{array}$$

$$c_0 = \pi_i+1 \qquad t = i.$$

REMARKS

In particular

$$c_1 = \frac{1}{\pi_j} \qquad \begin{array}{l} k_1 = j \\ k_1+t = i \quad \text{and} \quad \pi_i = 0. \end{array}$$

The terms on the right hand side of (7.11) are not necessarily distinct. When like terms are collected together all resulting coefficients are integers. As a numerical example

$$(2211)(1)=\tfrac{1}{2}(3211)+\tfrac{1}{2}(2311)+\tfrac{3}{2}(2221)+\tfrac{3}{2}(2212)+3(22111)$$
$$= (3211)+3(2221)+3(22111).$$

The table building algorithm is given by

COROLLARY

A rearrangement of (7.11) gives

$$(k_1\ldots k_r,t) = \frac{1}{c_0}\{(k_1\ldots k_r)(t)-c_1(k_1+t\ldots k_r)-\ldots\ldots$$
$$-c_r(k_1\ldots k_r+t)\}$$
$$(7.12)$$

REMARKS

This says that any $r+1$ part msf can be expressed in terms of r part msf's and a power sum. The power sums are easily calculated. For the 2 part msf's

$$(k)(t) = (k+t) + c_0(k,t)$$

where $c_0 = \begin{matrix} 1 & k \neq t \\ 2 & k = t \end{matrix}$

and hence

$$(k,t) = \frac{1}{c_0}\{(k)(t) - (k+t)\}.$$

Similarly we can calculate all required $r+1$ part msf's when the appropriate r part ones are known.

Note that if $k_i = k_{i+1}$ the two msf's $(k_1 \ldots k_i+t, k_{i+1} \ldots k_r)$ and $(k_1 \ldots k_i, k_{i+1}+t, \ldots k_r)$ are equal but are treated separately in (7.12) as in (7.11). This seems naive for hand methods but for computing purposes the "dumb" way is often the best way to make sure that all cases are considered. By hand it would be only necessary to determine the distinct msf's and multiply by the coefficients π_i+1.

The final step is to verify (7.11).

<u>Proof</u>

Let $(k_1\ k_2 \ldots k_r)$ be a msf in the indeterminates $s_1, s_2 \ldots s_m$.

Then

$$(k_1\ k_2 \ldots k_r)(t) = \{\Sigma\ s_u^{k_1} \ldots s_v^{k_q} \ldots s_w^{k_r}\}\{\sum_{v=1}^{m} s_v^{\,t}\}. \qquad (7.13)$$

Perform the product on the right hand side. Now the number of terms in which $s_v^{k_q+t}$ occurs, with $k_q = j$, is

$$\frac{[m-1]_{r-1}}{\pi_1!\ldots(\pi_j-1)!\ldots.} \tag{7.14}$$

However the indeterminate s_v can be associated with k_q in only $1/\pi_j$ of these terms as s_v must also be associated equally as often with the π_j-1 other parts also equal to j. Since v may take m possible values the total number of terms of the form $s_u^{k_1}\ldots s_v^{k_q+t}\ldots s_w^{k_r}$ is

$$\frac{m}{\pi_j}\,\frac{[m-1]_{r-1}}{\pi_1!\ldots(\pi_j-1)!\ldots} \quad . \tag{7.15}$$

Let $(k_1\ldots k_q+t\ldots k_r) = (1^{\pi_1}\ldots j^{\pi_j-1}\ldots i^{\pi_i+1}\ldots)$. Then the number of terms in its associated msf is

$$\frac{[m]_r}{\pi_1!\ldots(\pi_j-1)!\ldots(\pi_i+1)!\ldots} \tag{7.16}$$

and

$$c_q = \frac{(7.15)}{(7.16)} = \frac{\pi_i+1}{\pi_j} \; . \tag{7.17}$$

The term $s_u^{k_1}\ldots s_v^{k_r}s_w^{t}$ occurs in

$$\frac{m[m-1]_r}{\pi_1!\ldots\pi_i!\ldots} \tag{7.18}$$

ways, while if $(k_1\ldots.k_r, t) = (1^{\pi_1}\ldots i^{\pi_i+1}\ldots)$ this has

$$\frac{[m]_{r+1}}{\pi_1!\ldots(\pi_i+1)!\ldots} \tag{7.19}$$

terms and

$$c_0 = \frac{(7.18)}{(7.19)} = \pi_i+1 \, . \tag{7.20}$$

Q.E.D.

7.4 The formula for $\binom{\kappa}{\nu}$

The "binomial" coefficients $\binom{\kappa}{\nu}$ are defined by (6.12) while the $g^{\kappa}_{\nu\mu}$ are given by

$$C_\nu(S)\, C_\mu(S) = \sum_{\kappa} g^{\kappa}_{\nu\mu}\, C_\kappa(S) \tag{7.21}$$

where κ is a partition of $k = r+s$

ν is a partition of r

μ is a partition of s.

The relationship between $\binom{\kappa}{\nu}$ and $g^{\kappa}_{\nu\mu}$ is given by the

THEOREM 7.2

If k,r,s,κ,ν,μ are as defined above, then

$$\binom{\kappa}{\nu} = \binom{k}{s} \sum_{\mu} g^{\kappa}_{\nu\mu} \;. \tag{7.22}$$

The proof will follow from an easily established identity.

LEMMA 7.2

$${}_0F_0^{(m)}(I+A,B) = \mathrm{etr}(B)\,{}_0F_0^{(m)}(A,B). \tag{7.23}$$

Proof

From the integral definition the left hand side of (7.23) is

$${}_0F_0^{(m)}(I+A,B) = \int_{O(m)} \mathrm{etr}[(I+A)H'BH](dH)$$

$$= \mathrm{etr}(B) \int_{O(m)} \mathrm{etr}(AH'BH)(dH). \qquad \text{Q.E.D.}$$

Proof (THEOREM 7.2)

Writing the hypergeometric functions of (7.23) in zonal series

$$\sum_{k,\kappa} \frac{C_\kappa(I+A)C_\kappa(B)}{k!\ C_\kappa(I)} = \left\{\sum_{s,\mu} \frac{C_\mu(B)}{s!}\right\}\left\{\sum_{r,\nu} \frac{C_\nu(A)C_\nu(B)}{r!\ C_\nu(I)}\right\}. \quad (7.24)$$

Since all three zonal series converge everywhere in $A,B > 0$ we can take the Cauchy product of the right hand side to give

$$\sum_{k=0}^{\infty} \sum_{r+s=k} \frac{1}{r!s!} \sum_{\nu,\mu} \frac{C_\nu(A)}{C_\nu(I)} C_\nu(B)\ C_\mu(B)$$

$$\underset{(7.21)}{=} \sum_{k} \sum_{r+s=k} \frac{1}{r!s!} \sum_{\nu,\mu} \frac{C_\nu(A)}{C_\nu(I)} \sum_{\kappa} g^{\kappa}_{\nu\mu}\ C_\kappa(B).$$

Rearranging gives

$$\sum_{k,\kappa} \frac{C_\kappa(B)}{k!} \sum_{r+s=k} \sum_{\nu} \frac{C_\nu(A)}{C_\nu(I)} \binom{k}{s} \sum_{\mu} g^{\kappa}_{\nu\mu}.$$

Equating coefficients of $C_\kappa(B)/k!$ for this and the left hand side of (7.24) gives

$$\frac{C_\kappa(I+A)}{C_\kappa(I)} = \sum_{r=0}^{k} \sum_{\nu} \frac{C_\nu(A)}{C_\nu(I)} \binom{k}{s} \sum_{\mu} g^{\kappa}_{\nu\mu}. \quad (7.25)$$

The defining relation is

$$\frac{C_\kappa(I+A)}{C_\kappa(I)} = \sum_{r=0}^{k} \sum_{\nu} \frac{C_\nu(A)}{C_\nu(I)} \binom{\kappa}{\nu}$$

and the comparison of this with (7.25) gives (7.22). Q.E.D.

If we put $A = aI$ in the defining relation the result is

$$(1+a)^k = \sum_{r=0}^{k} a^r \sum_{\nu} \binom{\kappa}{\nu} \text{ for each } \kappa$$

and comparison with the binomial expansion of $(1+a)^k$ gives

LEMMA 7.3

$$\binom{k}{r} = \sum_{\nu} \binom{\kappa}{\nu} \quad \text{for all partitions } \kappa \text{ of } k. \tag{7.26}$$

This shows the connection between the "binomial" coefficients $\binom{\kappa}{\nu}$ and the true binomial coefficients $\binom{k}{r}$.

A computer program was written to determine the $\binom{\kappa}{\nu}$ for $k = 1,2....9$. The values obtained are in decimal form. The values $k=1,2,3,4$ have been given by CONSTANTINE [8] as fractions. The tables for $k = 5,6$ (see TABLES 7.3, 7.4) were obtained by printing out the values to 8 decimal places and converting these manually to fractions. It will be shown later on that the $g^{\kappa}_{\nu\mu}$ are rational and this implies that the $\binom{\kappa}{\nu}$ are too. The programming is discussed in Appendix 4.

7.5 The $g^{\kappa}_{\nu\mu}$ and products of monomial symmetric functions.

THEOREM 7.2 gives the relationship between $\binom{\kappa}{\nu}$ and $g^{\kappa}_{\nu\mu}$, but from a practical viewpoint this result is useless unless the $g^{\kappa}_{\nu\mu}$ are known. In this section is presented an algorithm for generating them. This method has been programmed for a computer evaluation of the $g^{\kappa}_{\nu\mu}$ and hence the $\binom{\kappa}{\nu}$ for κ a partition of k and $k = 1,2,...9$. Details are in Appendix 4.

Let $\kappa,\tau,\delta,\varepsilon$ be partitions of $k = r+t$
α,ν be partitions of r
β,μ be partitions of t.

k = 5

	κ: (5)	(41)	(32)	(31^2)	(2^21)	(21^3)	(1^5)
(0)	1	1	1	1	1	1	1
(1)	5	5	5	5	5	5	5
(2)	10	7	16/3	13/3	10/3	2	·
(1^2)	·	3	14/3	17/3	20/3	8	10
(3)	10	23/5	8/5	7/5	·	·	·
(21)	·	27/5	42/5	33/5	15/2	9/2	·
(1^3)	·	·	·	2	5/2	11/2	10
(4)	5	8/7	·	·	·	·	·
(31)	·	27/7	8/3	7/3	·	·	·
ν (2^2)	·	·	7/3	·	5/3	·	·
(21^2)	·	·	·	8/3	10/3	18/5	·
(1^4)	·	·	·	·	·	7/5	5
(5)	1	·	·	·	·	·	·
(41)	·	1	·	·	·	·	·
(32)	·	·	1	·	·	·	·
(31^2)	·	·	·	1	·	·	·
(2^21)	·	·	·	·	1	·	·
(21^3)	·	·	·	·	·	1	·
(1^5)	·	·	·	·	·	·	1

TABLE 7.3 The $\binom{\kappa}{\nu}$

k = 6

	κ → (6)	(51)	(42)	(41^2)	(3^2)	(321)	(31^3)	(2^3)	(2^21^2)	(21^4)	(1^6)
ν ↓ (0)	1	1	1	1	1	1	1	1	1	1	1
(1)	6	6	6	6	6	6	6	6	6	6	6
(2)	15	34/3	9	8	8	19/3	5	5	4	7/3	·
(1^2)	·	11/3	6	7	7	26/3	10	10	11	38/3	15
(3)	20	56/5	28/5	26/5	16/5	28/15	8/5	·	·	·	·
(21)	·	44/5	72/5	123/10	84/5	74/5	57/5	15	12	7	·
(1^3)	·	·	·	5/2	·	10/3	7	5	8	13	20
(4)	15	39/7	48/35	9/7	·	·	·	·	·	·	·
(31)	·	66/7	66/7	61/7	8	14/3	4	·	·	·	·
(2^2)	·	·	21/5	·	7	11/3	·	5	5/2	·	·
(21^2)	·	·	·	5	·	20/3	46/5	10	52/5	42/5	·
(1^4)	·	·	·	·	·	·	9/5	·	21/10	33/5	15
(5)	6	10/9	·	·	·	·	·	·	·	·	·
(41)	·	44/9	12/5	9/4	·	·	·	·	·	·	·
(32)	·	·	18/5	·	6	14/9	·	·	·	·	·
(31^2)	·	·	·	15/4	·	20/9	24/7	·	·	·	·
(2^21)	·	·	·	·	·	20/9	·	6	3	·	·
(21^3)	·	·	·	·	·	·	18/7	·	3	14/3	·
(1^5)	·	·	·	·	·	·	·	·	·	4/3	6
(6)	1	·	·	·	·	·	·	·	·	·	·
(51)	·	1	·	·	·	·	·	·	·	·	·
(42)	·	·	1	·	·	·	·	·	·	·	·
(41^2)	·	·	·	1	·	·	·	·	·	·	·
(3^2)	·	·	·	·	1	·	·	·	·	·	·
(321)	·	·	·	·	·	1	·	·	·	·	·
(31^3)	·	·	·	·	·	·	1	·	·	·	·
(2^3)	·	·	·	·	·	·	·	1	·	·	·
(2^21^2)	·	·	·	·	·	·	·	·	1	·	·
(21^4)	·	·	·	·	·	·	·	·	·	1	·
(1^6)	·	·	·	·	·	·	·	·	·	·	1

TABLE 7.4 The $\binom{\kappa}{\nu}$

Assume for convenience $r \geqslant t$. Drop the argument matrix i.e. put $C_\kappa = C_\kappa(\cdot)$, $M_\kappa = M_\kappa(\cdot)$ as the matrix or its size does not enter into any of the following relationships.

From (7.4)

$$C_\kappa = \sum_{\tau \leqslant \kappa} a_{\kappa\tau} M_\tau \tag{7.27}$$

where $a_{\kappa\tau} = c(\kappa)\, c_{\kappa\tau}$. Now using (7.27)

$$C_\nu C_\mu = \sum_{\alpha \leqslant \nu} \sum_{\beta \leqslant \mu} a_{\nu\alpha} a_{\mu\beta} M_\alpha M_\beta . \tag{7.28}$$

Let $M_\alpha = (f_1 \ldots f_r)$ $M_\beta = (h_1 \ldots h_t)$ and it is essential in the following that none of the parts f_i and h_j are zero. Then the msf of highest weight obtainable from the product of M_α and M_β is $M_\tau = (f_1+h_1 \ldots f_t+h_t \ldots f_r)$. Let

$$M_\alpha M_\beta = \sum_{\delta \leqslant \tau} e_{\alpha\beta}^{\delta} M_\delta \tag{7.29}$$

(analogous to (7.21) for $g_{\nu\mu}$). Then (7.28) can be written as

$$C_\nu C_\mu = \sum_{\alpha \leqslant \nu} \sum_{\beta \leqslant \mu} \sum_{\delta \leqslant \tau} a_{\nu\alpha} a_{\mu\beta} e_{\alpha\beta}^{\delta} M_\delta . \tag{7.30}$$

The msf's of highest weight in C_ν and C_μ are M_ν and M_μ respectively. Let the msf of highest weight in $M_\nu M_\mu$ be M_κ, then if we define $e_{\alpha\beta}^{\delta} = 0$ $\tau < \delta \leqslant \kappa$, (7.30) can be written as

$$C_\nu C_\mu = \sum_{\delta \leqslant \kappa} d_{\nu\mu}^{\delta} M_\delta \tag{7.31}$$

where

$$d_{\nu\mu}^{\delta} = \sum_{\alpha \leqslant \nu} \sum_{\beta \leqslant \mu} a_{\nu\alpha} a_{\mu\beta} e_{\alpha\beta}^{\delta} .$$

The defining relation for the $g^{\kappa}_{\nu\mu}$ can be written as

$$C_\nu \; C_\mu = \sum_{\delta \leqslant \kappa} g^{\delta}_{\nu\mu} \; C_\delta \tag{7.32}$$

so a sequential comparison of the coefficients in (7.31) and (7.32) will give us the $g^{\delta}_{\nu\mu}$.

For example

$$C_\kappa = \sum_{\varepsilon \leqslant \kappa} a_{\kappa\varepsilon} \; M_\varepsilon$$

and equating the coefficients of M_κ gives

$$g^{\kappa}_{\nu\mu} = \frac{d^{\kappa}_{\nu\mu}}{a_{\kappa\kappa}}. \tag{7.33}$$

Subtract $g^{\kappa}_{\nu\mu} \; C_\kappa$ from (7.31)

$$d^{\kappa}_{\nu\mu} \to 0$$

$$d^{\delta}_{\nu\mu} \to d^{\delta}_{\nu\mu} - g^{\kappa}_{\nu\mu} \; a_{\kappa\delta} \qquad\qquad \delta < \kappa.$$

Then if τ is immediately below κ in the lexicographic list

$$g^{\tau}_{\nu\mu} = \frac{d^{\tau}_{\nu\mu}}{a_{\tau\tau}}$$

and so on. Thus we can express the $g^{\kappa}_{\nu\mu}$ in terms of the $a_{\kappa\delta}$ (known) and the $e^{\delta}_{\alpha\beta}$.

The calculation of the $e^{\delta}_{\alpha\beta}$ is quite straightforward, at least in principle. Let

$$M_\alpha \; M_\beta = \{\Sigma \; s_i^{f_1} s_j^{f_2} \ldots s_q^{f_r}\}\{\Sigma \; s_u^{h_1} s_v^{h_2} \ldots . s_w^{h_t}\} \tag{7.34}$$

and take all possible products i.e.

$$M_\alpha \; M_\beta = \sum_{\varepsilon \leqslant \tau} b_\varepsilon \; M_\varepsilon \tag{7.35}$$

where the ε are not necessarily distinct partitions. Adding together all b_ε for the same partition will give the $e^{\delta}_{\alpha\beta}$.

What are the b_ε?

THEOREM 7.3

Let $\alpha = (f_1 \ldots f_r) = (1^{\varphi_1}2^{\varphi_2}\ldots)$

$\beta = (h_1 \ldots h_t) = (1^{\psi_1}2^{\psi_2}\ldots)$

$\varepsilon = (\ell_1 \ldots \ell_n) = (1^{\lambda_1}2^{\lambda_2}\ldots)$

where M_ε is a possible product, then

$$b_\varepsilon = \frac{\lambda_1!\lambda_2!\ldots}{\varphi_1!\varphi_2!\ldots\psi_1!\psi_2!\ldots} . \qquad (7.36)$$

REMARK

As an example

$$\begin{aligned}(211)(11) \underset{(7.35)}{=} & \tfrac{1}{4}(321) + \tfrac{1}{4}(321) + \tfrac{1}{4}(312) + \tfrac{1}{4}(312) \\ & + \tfrac{6}{4}(222) + \tfrac{6}{4}(222) + \tfrac{6}{4}(3111) + \tfrac{4}{4}(2211) \\ & + \tfrac{4}{4}(2121) + \tfrac{6}{4}(3111) + \tfrac{4}{4}(2211) + \tfrac{4}{4}(2121) + \tfrac{24}{4}(21111) \\ \underset{(7.29)}{=} & (321) + 3(3111) + 3(222) + 4(2211) + 6(21111).\end{aligned}$$

<u>Proof</u>

finally a proof of (7.36) when the product term of (7.34) has the form

$$s_1^{f_1} \ldots s_q^{f_q+h_u} \ldots s_v^{f_v+h_w} \ldots s_r^{f_r} s_x^{h_1} \ldots s_z^{h_t} .$$

There are two cases

1. $f_q = f_v = i$.

With $q, v, x \ldots z$ fixed, the number of possible terms is

$$\frac{[m-t]_{r-2}}{\varphi_1!\ldots(\varphi_i-2)!\ldots} .$$

Now s_q, s_v can be associated with f_q, f_v in $\frac{1}{\varphi_i(\varphi_i-1)}$ terms and the number of ways the indices $q,v,x\ldots z$ can be chosen is

$$\frac{[m]_t}{\psi_1!\psi_2!\ldots} . \tag{7.37}$$

Thus the total number of terms is

$$\frac{[m]_{r+t-2}}{\varphi_1!\varphi_2!\ldots\psi_1!\psi_2!\ldots} . \tag{7.38}$$

Let $\varepsilon = (f_1\ldots f_q+h_u\ldots f_v+h_w\ldots f_r, h_1\ldots h_t) = (1^{\lambda_1}2^{\lambda_2}\ldots)$ and the associated msf has total number of terms

$$\frac{[m]_{r+t-2}}{\lambda_1!\lambda_2!\ldots} . \tag{7.39}$$

Then $$b_\varepsilon = \frac{(7.38)}{(7.39)}$$ as required.

2. $f_q = i \quad f_v = j$.

With $q,v,x,\ldots z$ fixed, the number of terms is

$$\frac{[m-t]_{r-2}}{\varphi_1!\ldots(\varphi_i-1)!\ldots(\varphi_j-1)!\ldots} .$$

Now s_q can be associated with f_q in $\frac{1}{\varphi_i}$ ways,

s_v can be associated with f_v in $\frac{1}{\varphi_j}$ ways,

and the number of ways $q,v,x\ldots z$ can be chosen is still given by (7.37). The rest of the proof is the same as for case 1.

In the particular case $t=1$, the formulae agree with those for c_q and c_0.

The coefficients $g^{\kappa}_{\nu\mu}$ have already been tabled for $k = 1,2,\ldots 7$ by KHATRI and PILLAI [25]. Their methods were based on the expansion of the zonal polynomials in terms of the elementary symmetric functions and the power sums. The limiting factor in this approach is the comparative difficulty of expressing the zonal polynomials in terms of these functions versus their expression in msf's.

7.6 Summation identities for $\binom{\kappa}{\nu}$ and $g^{\kappa}_{\nu\mu}$.

This chapter is concluded with some formulae that are useful in checking that the values for $\binom{\kappa}{\nu}$ and $g^{\kappa}_{\nu\mu}$ are correct. One such formula (7.26) has already been given. All results are obtained by using identities similar to (7.23) and multiplying out. Coefficients of $C_{\kappa}(S)$ are then equated. Throughout this section k,r,s,κ,ν,μ are as defined for (7.21). The matrix S is $m\times m$ positive definite.

IDENTITY 1

$$\operatorname{etr}[(x+y)S] = \operatorname{etr}(xS)\operatorname{etr}(yS) \qquad (7.40)$$

LEMMA 7.4

$$(x+y)^k = \sum_{r+s=k} x^r y^s \binom{k}{r}_{\nu} \sum_{,\mu} g^{\kappa}_{\nu\mu} \qquad (7.41)$$

Proof

Expand the term on the left hand side of (7.40) in a zonal series. Expand the terms on the right hand side similarly and take their Cauchy product. Equate the coefficients of $C_\kappa(S)/k!$. Q.E.D.

COROLLARY 1

$$1 = \sum_{\nu,\mu} g^{\kappa}_{\nu\mu} \quad \text{for } r,s \text{ fixed.} \tag{7.42}$$

Proof

Set $y=1$ in (7.41).

$$(1+x)^k = \sum_{r=0}^{k} \binom{k}{r} x^r = \sum_{r=0}^{k} \binom{k}{r} x^r \sum_{\nu,\mu} g^{\kappa}_{\nu\mu} . \qquad \text{Q.E.D.}$$

COROLLARY 2

$$2^k = \sum_{r+s=k} \binom{k}{r} \sum_{\nu,\mu} g^{\kappa}_{\nu\mu} . \tag{7.43}$$

Proof

Set $x=y=1$ in (7.41). Q.E.D.

Many other formulae can be easily established. These are perhaps mainly useful for checking the tables of $g^{\kappa}_{\nu\mu}$.

IDENTITY 2

$$\det(I-S)^{-a-b} = \det(I-S)^{-a} \det(I-S)^{-b} \qquad a,b \text{ real} \tag{7.44}$$

or in hypergeometric function notation

$${}_1F_0(a+b;S) = {}_1F_0(a;S)\,{}_1F_0(b;S). \tag{7.45}$$

LEMMA 7.5

$$(a+b)_\kappa = \sum_{r+s=k} \binom{k}{r} \sum_{\nu,\mu} (a)_\nu (b)_\mu \, g^\kappa_{\nu\mu}. \tag{7.46}$$

Proof

Expand both sides of (7.45) in zonal series, perform the Cauchy product and equate coefficients. Q.E.D.

For fixed κ, various values of a and b may be chosen to generate a set of simultaneous linear equations for the $g^\kappa_{\nu\mu}$. All coefficients can be chosen as rational by taking a,b rational. Thus the $g^\kappa_{\nu\mu}$ are rational.

IDENTITY 3

JAMES [21] lists the KUMMER relation (equation (51))

$$_1F_1(a;b;S) = \mathrm{etr}(S)\,_1F_1(b-a;b;-S). \tag{7.47}$$

LEMMA 7.6

$$\frac{(a)_\kappa}{(b)_\kappa} = \sum_{s=0}^{k} (-1)^s \binom{k}{s} \sum_{\nu,\mu} \frac{(b-a)_\mu}{(b)_\mu} \, g^\kappa_{\nu\mu}. \tag{7.48}$$

Proof

Expand in zonal series, etc. Q.E.D.

COROLLARY

$$\frac{(a)_\kappa}{(b)_\kappa} = \sum_{s=0}^{k} (-1)^s \sum_{\mu} \frac{(b-a)_\mu}{(b)_\mu} \binom{\kappa}{\mu}. \tag{7.49}$$

Proof

Rearrange (7.48) as

$$\frac{(a)_\kappa}{(b)_\kappa} = \sum_{s=0}^{k} (-1)^s \sum_{\mu} \frac{(b-a)_\mu}{(b)_\mu} \binom{k}{r} \sum_{\nu} g^\kappa_{\nu\mu}$$

and apply (7.22). Q.E.D.

Both (7.48) and (7.49) could be used to give systems of simultaneous linear equations for $g^\kappa_{\nu\mu}$ and $\binom{\kappa}{\mu}$.

CHAPTER 8

NUMERICAL EVALUATION

8.1 Introduction

It now remains to consider the arithmetic worth of the various formulae for the numerical calculation of the one and two argument Bessel functions. The evaluations are over three ranges, one each for small, medium and large values of the latent roots. A section is devoted to each range of values and within each section both the one and two argument functions are considered.

Each section begins with an outline of the formulae used and this is followed by the results obtained when a few specific values are input to computer programmes written to perform the evaluation.

Results are good for very small and very large values. The limited results obtained for some medium values are encouraging but inconclusive in the case of the single Laguerre expansion. Results for the double Laguerre expansion are rather discouraging. However an extensive computer evaluation programme would be needed to verify these assertions.

The value $n=10$ was used in all evaluations.

8.2 Small latent roots - zonal series

Direct summation of the zonal series is the method to be used when the latent roots are all small. Both the one and two argument Bessel functions can be evaluated using the same computer programme by making use of the identity (1.17). That is

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;R,S) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(R)C_\kappa(S)}{(\frac{1}{2}n)_\kappa k! C_\kappa(I)} \tag{8.1} $$

and

$$ {}_0F_1(\tfrac{1}{2}n;R) = {}_0F_1^{(m)}(\tfrac{1}{2}n;R,I). \tag{8.2} $$

It was decided to restrict the evaluations to the cases m=2 and m=3.

One argument matrix:

	R			value	sig. figs.
m=2		2	1	1·7961424950	11
		4	2	3·14383941	9
		8	4	9·034	4
		16	8	61·4	3
m=3	3	2	1	3·202536	7
	6	4	2	9·624	4
	12	8	4	74·5	3

Two argument matrices:

	R			S			value	sig. figs.
m=2		2	1		2	1	2·38036499	9
		4	2		4	2	23·94	4
		8	4		8	4	2×10^4	1
		16	8		16	8	-	none
m=3	3	2	1	3	2	1	9·5360	5
	6	4	2	6	4	2	250	2
	12	8	4	12	8	4	-	none

It is clear that very good accuracy is obtainable for latent roots with values less than 1. For one argument matrix about 4 significant figure accuracy is obtainable if the leading latent root is less than about 8 while for two argument matrices, to obtain the same accuracy the leading latent root should be less than 4 (only for n=10).

8.3 Medium value latent roots - Laguerre series

First the single Laguerre series. Substituting for $L_\kappa^a(S)$ using (6.11), noting that $p = \frac{1}{2}(m+1)$ and using the scaled zonal polynomials $C_\kappa{}^*(S)$ defined by

$$C_\kappa{}^*(S) = \frac{C_\kappa(S)}{C_\kappa(I)} = \frac{Z_\kappa(S)}{Z_\kappa(I)} \tag{8.3}$$

the equation (6.19) becomes

$$ {}_0F_1^{(m)}(a+p;S,Z) = $$

$$\operatorname{etr}(Z) \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sum_\kappa C_\kappa(Z) \sum_{n=0}^{k} (-1)^n \sum_\nu \binom{\kappa}{\nu} \frac{C_\nu{}^*(S)}{(a+p)_\nu}. \tag{8.4}$$

Setting $a+p = \frac{1}{2}n$, this is evaluated in the form

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;S,Z) = \operatorname{etr}(Z) \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sum_\kappa C_\kappa(Z) A_\kappa(S) \tag{8.5}$$

where

$$A_\kappa(S) = \sum_{n=0}^{k} (-1)^n \sum_\nu \binom{\kappa}{\nu} \frac{C_\nu{}^*(S)}{(\frac{1}{2}n)_\nu}. \tag{8.6}$$

As a test, values were chosen for which the zonal series converged to an answer after summing all terms to k=9. For n=10, m=2, S=diag(7·5,2·5). Z=diag(5,0·5) a value of 108 (correct to 3 significant figures) was obtained. Using these values in the Laguerre series (8.5) no satisfactory results were obtained while summing to $k=9$.

By introducing a scaling factor $c \neq 0$ considerable improvement is possible. We have the identity

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;cR,c^{-1}S) = {}_0F_1^{(m)}(\tfrac{1}{2}n;R,S) \tag{8.7}$$

which follows easily from the fact that

$$C_\kappa(cR) = c^k C_\kappa(R). \tag{8.8}$$

Thus from (8.7), (8.5) becomes

$$\begin{aligned} {}_0F_1^{(m)}(\tfrac{1}{2}n;S,Z) &= {}_0F_1^{(m)}(\tfrac{1}{2}n;cS,c^{-1}Z) \\ &= \operatorname{etr}(c^{-1}Z) \sum_k \frac{(-1)^k}{k!} \sum_\kappa C_\kappa(Z)A_\kappa(cS)c^{-k} \end{aligned} \tag{8.9}$$

and clearly from (8.6)

$$A_\kappa(cS) \neq c^k A_\kappa(S).$$

Summing the series (8.9) using the above values and $c=10$ also gave 108. This indicates that it may be possible to obtain improved convergence by a suitable choice of a scaling factor. This introduction of an extra parameter is not possible in (8.1) as

$$C_\kappa(cR)C_\kappa(c^{-1}S) = C_\kappa(R)C_\kappa(S)$$

and the effect is lost unlike the case (8.9).

Similar results can be obtained for the one matrix case by setting

$$ {}_0F_1(\tfrac{1}{2}n;R) = {}_0F_1^{(m)}(\tfrac{1}{2}n;cR,c^{-1}I). \tag{8.10}$$

As already stated no sensible results were obtained for the double Laguerre series even with the introduction of a scale factor. A far more extensive study would be necessary however to verify this as well as to confirm and perhaps improve on results obtained from (8.9).

8.4 Large latent roots - asymptotic formulae

The asymptotic formula used for the one argument matrix case is, from (4.22) and (4.19)

$$ {}_0F_1(\tfrac{1}{2}n;\tfrac{1}{4}XX') = \frac{k_1 \operatorname{etr}(A)}{\prod_{i<j}(a_i+a_j)^{\frac{1}{2}} \det A^{\frac{1}{2}(n-m)}} G(A) \qquad (8.11) $$

where k_1 is given in (4.22) and

$$ G(A) = 1 - \Lambda_1 \sum_{i=1}^{m} \frac{1}{a_i} - \Lambda_2 \sum_{i<j} \frac{1}{a_i+a_j} + o\left(\frac{1}{a}\right) \qquad (8.12) $$

where

$$ \Lambda_1 = \tfrac{1}{8}(n-3)(n-m) \qquad \Lambda_2 = \tfrac{1}{12}(m-2). $$

For two argument matrices, THEOREM 2.4 gives

$$ {}_0F_1^{(m)}(\tfrac{1}{2}n;\tfrac{1}{4}\Omega,W) = k_2 \frac{\operatorname{etr}(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}} \det(AB)^{\frac{1}{2}(n-m)}} G(A,B) \qquad (8.13) $$

where k_2 is given in (2.58) and

$$ G(A,B) = 1 - \Lambda_1 \sum_{i=1}^{m} \frac{1}{a_i b_i} - 2\Lambda_2 \sum_{i\neq j} \frac{a_i b_i}{c_{ij}} + o\left(\frac{1}{a^2}\right). \qquad (8.14) $$

A simple lower bound on the values of a_i for (8.11) and of a_i, b_i for (8.13) can be obtained simply. For two matrices first. Assuming the a_i, b_i are large and well spaced, we have

$$c_{ij} = (a_i^2 - a_j^2)(b_i^2 - b_j^2)$$
$$= a_i^2 b_i^2\left(1 - \frac{a_j^2}{a_i^2}\right)\left(1 - \frac{b_j^2}{b_i^2}\right)$$
$$\simeq a_i^2 b_i^2$$

and

$$\frac{a_i b_i}{c_{ij}} \simeq \frac{1}{a_i b_i} . \qquad (8.15)$$

Also since the a_i and b_i are ordered in a decreasing sequence

$$\frac{1}{a_m b_m} \geqslant \frac{1}{a_i b_i} \qquad i = 1,2,\ldots m. \qquad (8.16)$$

Hence since the Bessel function must be positive we would like to have

$$\Lambda_1 \sum_{i=1}^{m} \frac{1}{a_i b_i} + 2\Lambda_2 \sum_{i \neq j} \frac{a_i b_i}{c_{ij}} < 1$$

and using (8.15) and (8.16) this becomes

$$\frac{1}{a_m b_m}(\Lambda_1 m + 2\Lambda_2 m(m-1)) < 1.$$

The solution is

$$a_m b_m > \tfrac{1}{8}(n-3)(n-m)m + \tfrac{1}{6}m(m-1)(m-2). \qquad (8.17)$$

For $m=2$, $n=10$ this gives

$$a_m b_m > 14, \quad \text{i.e.} \quad a_m, b_m = 4$$

and for $m=3$, $n=10$

$$a_m b_m > 19, \quad \text{i.e.} \quad a_m = 5,\ b_m = 4.$$

A programme was written to calculate the first four correction terms for $G(A,B) \simeq 1-T_1-T_2+T_3+T_4$.

	a_i	b_i	T_1	T_2	T_3	T_4
m=2	8, 4	8, 4	0.54	0	0	.09
	16, 8	16, 8	0.13	0	0	.005
	24,12	24,12	0.009	0	0	.0025
m=3	15,10, 5	12, 8, 4	0.41	0.01	.004	.046

Thus (8.17) provides an adequate lower bound when the higher values have the form $a_i = (m-i+1)a_m \quad i = 1,2,\ldots m-1$.

Also for one argument matrix we arrive at the bound

$$a_m > \tfrac{1}{8}(n-3)(n-m)m + \tfrac{1}{24}m(m-1)(m-2). \qquad (8.18)$$

For $m=2$, $n=10$, $a_m > 14$ and for $m=3$, $n=10$, $a_m > 19$.

Similarly writing $G(A) \simeq 1-T_1-T_2+T_3+T_4$.

	a_i	T_1	T_2	T_3	T_4
m=2	30,15	0.7	0	0.12	0.14
	60,30	0.35	0	.03	.03
	90,30	0.31	0	.03	.02
m=3	60,40,20	0.56	.003	.06	.12
	90,60,30	0.37	.002	.03	.05
	120,80,40	0.28	.001	.01	.03

Thus (8.18) also provides a lower bound on the a_i in the sense that if a_m does not satisfy it, absurd values are certain to result.

In most cases $1-T_1-T_2$ will lead to an approximation with at least 2 significant figures, however much higher values of the latent roots are necessary for the leading terms themselves to provide an adequate approximation to the true value.

8.5 Concluding remarks

Excellent results are obtained for small and large values of the latent roots. It appears though that much work is necessary to produce conclusive results on the worth of introducing a scale factor in the single Laguerre expansion and on the double Laguerre expansion.

Another interesting possibility is that the asymptotic series of Chapter 5 may produce better results for small values of the latent roots. The removal of the term etr(2S) may well yield more rapid convergence than the zonal series. Similarly, the Laguerre series too may very well be more rapidly convergent for certain small values of the latent roots. Certainly much theoretical and numerical work remains to be done.

APPENDIX 1

A.1.1 Calculation of the d_κ

The coefficients d_κ are related to the c_κ by (2.20) and these can be evaluated from (2.17). From the formula of CONSTANTINE [7] equation (31)

$$\det(I-\tfrac{1}{2}U)^{\frac{1}{2}(n-2m-1)} = {}_1F_0(-\tfrac{1}{2}(n-2m-1);\tfrac{1}{2}U) \qquad (A.1.1)$$

$$= \sum_{k,\kappa} (-\tfrac{1}{2}(n-2m-1))_\kappa \frac{C_\kappa(\frac{1}{2}U)}{k!} .$$

Also $\prod_{i<j} (1 - \frac{u_i+u_j}{2})$ is a symmetric function of the $u_1,\ldots u_m$.

Let $a_i{}^*$ be the i^{th} elementary symmetric function of the u_i, then the product has the expansion

$$\prod_{i<j} (1 - \frac{u_i+u_j}{2}) = 1 + \gamma_1 a_1{}^* + \gamma_2 a_2{}^* + \gamma_3 a_1{}^{*2} + o(u^2).$$

To evaluate the γ_i, consider the product in an array form as

$$\left(1 - \frac{u_1+u_2}{2}\right)\left(1 - \frac{u_1+u_3}{2}\right)\left(1 - \frac{u_1+u_4}{2}\right) \cdots \left(1 - \frac{u_1+u_m}{2}\right)$$

$$\left(1 - \frac{u_2+u_3}{2}\right)\left(1 - \frac{u_2+u_4}{2}\right) \cdots \left(1 - \frac{u_2+u_m}{2}\right)$$

$$\vdots$$

$$\left(1 - \frac{u_{m-1}+u_m}{2}\right).$$

By the symmetry it is sufficient to count the number of times a typical term occurs in order to determine the γ_i.

To obtain a term of degree r, choose u_1's from r terms and 1's from the remainder.

We will find the coefficients of u_1, u_1^2 and $u_1 u_2$.

For degree 1.

Note that u_1 occurs in terms of row 1 only and there are $m-1$ terms. The coefficient of $-\frac{1}{2}u_1$ is $m-1$.

For degree 2.

Similarly u_1^2 can be obtained from the terms of row 1 in $\binom{m-1}{2}$ ways.

The product u_1u_2 is obtained in two ways. Choose u_2 from the leading term and u_1 from any other term in row 1. This can be done in $m-2$ ways. Also choose u_1 from any term in row 1 and u_2 from any term in row 2. This can be done in $(m-1)(m-2)$ ways. Combining, the coefficient of $\frac{1}{4}u_1u_2$ is $m(m-2)$.

To express the results in terms of the a_i^* remember that

$$\Sigma u_i^2 = a_1^{*2} - 2a_2^*$$

and the expansion from the array is

$$1 - \tfrac{1}{2}(m-1)a_1^* + \tfrac{1}{8}(m-1)(m-2)a_1^{*2} + \tfrac{1}{4}(m-2)a_2^* + o(u^2). \qquad (A.1.2)$$

JAMES [21] gives tables of the zonal polynomials in terms of the elementary symmetric functions and these can be solved to express the a_i^* in terms of the $C_\kappa(U)$

$$a_1^* = C_{(1)}(U) \qquad a_1^{*2} = C_{(2)}(U) + C_{(1^2)}(U)$$
$$a_2^* = \tfrac{3}{4}C_{(1^2)}(U).$$

Substituting in (A.1.2) and collecting terms gives

$$1 - \tfrac{1}{2}(m-1)C_{(1)}(U) + \tfrac{1}{8}(m-1)(m-2)C_{(2)}(U) + \tfrac{1}{16}(m-2)(2m+1)C_{(1^2)}(U) + o(u^2). \qquad (A.1.3)$$

Multiplying the two series (A.1.1) and (A.1.3) and using the zonal product formula

$$[C_{(1)}(U)]^2 = a_1^{*2} = C_{(2)}(U) + C_{(1^2)}(U)$$

we get

$$1 - \tfrac{1}{4}(n-3)C_{(1)}(U) + \tfrac{1}{32}(n-3)(n-5)C_{(2)}(U) + \tfrac{1}{32}n(n-5)C_{(1^2)}(U) + o(u^2). \qquad (A.1.4)$$

As a check none of the c_κ coefficients depend on m.

A.1.2 <u>The evaluation of (2.42)</u>

There are three types of integral involved. All can be evaluated using standard bivariate normal integrals.

LEMMA A.1.1

Let P, Q be 2×2 symmetric matrices and $\underline{s} = (s_1, s_2)'$. Then if Q is positive definite

$$\int_{-\infty}^{\infty}\!\!\int \exp(-\tfrac{1}{2}\underline{s}'Q\underline{s})\,d\underline{s} = \frac{2\pi}{\det Q^{\frac{1}{2}}} \qquad (A.1.5)$$

$$\int_{-\infty}^{\infty}\!\!\int \underline{s'Ps}\exp(-\tfrac{1}{2}\underline{s}'Q\underline{s})\,d\underline{s} = \frac{2\pi}{\det Q^{\frac{1}{2}}}\operatorname{tr}(Q^{-1}P) \qquad (A.1.6)$$

$$\int_{-\infty}^{\infty}\!\!\int \underline{s}'\underline{s}\exp(-\tfrac{1}{2}\underline{s}'Q\underline{s})\,d\underline{s} = \frac{2\pi}{\det Q^{\frac{1}{2}}}\operatorname{tr}(Q^{-1}). \qquad (A.1.7)$$

((A.1.7) is of course the particular case of (A.1.6) with $P = I_2$.)

Considering (2.42) we have

$$c_{ij} = \det Q_{ij} = (a_i^2 - a_j^2)(b_i^2 - b_j^2) > 0$$

and the leading principal minor is $a_i b_i + a_j b_j > 0$, hence Q_{ij} is positive definite. Thus

$$K(A,B) = \prod_{i<j} \int_{-\infty}^{\infty}\int \exp(-\tfrac{1}{2}\underline{s}_{ij}' Q_{ij} \underline{s}_{ij}) d\underline{s}_{ij} \qquad (A.1.8)$$

$$\underset{(A.1.5)}{=} \frac{(2\pi)^{\frac{1}{2}m(m-1)}}{\prod_{i<j} c_{ij}^{\frac{1}{2}}}.$$

Also to calculate the terms required for (2.47)

$$(1+\varphi(S,T;A,B))J(S)J(T) = (1-d_1(\sum_i \alpha_i \beta_i - \tfrac{1}{2}\sum_{i<j} \underline{s}_{ij}' P_{ij} \underline{s}_{ij} + o(s^2)) + ..)$$
$$\times(1 - \tfrac{1}{12}(m-2)\sum_{i<j} \underline{s}_{ij}' \underline{s}_{ij} + o(s^2))$$
$$(A.1.9)$$

where $d_1 = \frac{1}{8}(n-3)(n-m)$. The substitution of (A.1.9) in (2.40) gives

$$g(A,B) = \operatorname{etr}(AB)\{K(A,B) - d_1 \Sigma\alpha_i\beta_i K(A,B) + \tfrac{1}{2} d_1 L_1(A,B)$$
$$- \tfrac{1}{12}(m-2)L_2(A,B) + \tfrac{1}{12}(m-2) d_1 \Sigma \alpha_i \beta_i L_2(A,B) + \ldots\}$$
$$(A.1.10)$$

where

$$L_1(A,B) = \int_{\mathcal{S}}\int_{\mathcal{T}} \sum_{i<j} \underline{s}_{ij}' P_{ij} \underline{s}_{ij} \exp(-\tfrac{1}{2}\sum_{i<j} \underline{s}_{ij}' Q_{ij} \underline{s}_{ij}) d\underline{s}_{ij} \qquad (A.1.11)$$

$$\underset{(A.1.6)}{=} \frac{(2\pi)^{\frac{1}{2}m(m-1)}}{\prod_{i<j} c_{ij}^{\frac{1}{2}}} \sum_{i<j} \operatorname{tr}(Q_{ij}^{-1} P_{ij})$$

and

$$L_2(A,B) = \int_{\mathcal{S}}\int_{\mathcal{J}} \sum_{i<j} \underline{s}_{ij}'\underline{s}_{ij} \exp(-\tfrac{1}{2}\sum_{i<j}\underline{s}_{ij}'Q_{ij}\underline{s}_{ij})\, d\underline{s}_{ij} \qquad (A.1.12)$$

$$\underset{(A.1.7)}{=} \frac{(2\pi)^{\frac{1}{2}m(m-1)}}{\prod_{i<j} c_{ij}^{\frac{1}{2}}} \sum_{i<j} tr(Q_{ij}^{-1}).$$

Now $P_{ij} = \alpha_i\alpha_j\beta_i\beta_j Q_{ij}$ so $tr(Q_{ij}^{-1}P_{ij}) = 2\alpha_i\alpha_j\beta_i\beta_j$

and

$$Q_{ij}^{-1} = \frac{1}{c_{ij}} \begin{bmatrix} a_ib_i+a_jb_j & -a_ib_j-a_jb_i \\ -a_ib_j-a_jb_i & a_ib_i+a_jb_j \end{bmatrix}$$

giving $\quad tr(Q_{ij}^{-1}) = \dfrac{2(a_ib_i+a_jb_j)}{c_{ij}}$.

Substitution of these results in L_1 and L_2 and their substitution in (A.1.10) gives

$$g(A,B) = \frac{(2\pi)^{\frac{1}{2}m(m-1)} etr(AB)}{\prod_{i<j} c_{ij}^{\frac{1}{2}}} \Big\{ 1 - \tfrac{1}{8}(n-3)(n-m)\sum_{i=1}^{m}\frac{1}{a_ib_i}$$

$$+ \tfrac{1}{16}(n-3)(n-m)\sum_{i<j}\frac{1}{a_ib_ia_jb_j} - \tfrac{1}{6}(m-2)\sum_{i<j}\frac{a_ib_i+a_jb_j}{c_{ij}}$$

$$+ \tfrac{1}{48}(n-3)(n-m)(m-2)\sum_{k=1}^{m}\frac{1}{a_kb_k}\sum_{i<j}\frac{a_ib_i+a_jb_j}{c_{ij}} + \dots \Big\}.$$

(A.1.13)

To list the terms in increasing powers of $\frac{1}{a_ib_i}$ note that

$$\frac{1}{c_{ij}} = \frac{1}{a_i^2b_i^2\left(1 - \frac{a_j^2}{a_i^2}\right)\left(1 - \frac{b_j^2}{b_i^2}\right)} > \frac{1}{a_i^2b_i^2} .$$

Some rearrangements of terms then give the results (2.46) and (2.47).

A.1.3 Further terms of the series

Putting $R^{-1} = B^{-1}H_1'A^{-1}H_2'$ the next terms of $F(H_1,H_2;A,B)$ are $d_{(2)}C_{(2)}(R^{-1})$ and $d_{(1^2)}C_{(1^2)}(R^{-1})$. Expressing the zonal polynomials in terms of elementary symmetric functions gives

$$\tfrac{1}{128}(n-3)(n-5)(n-m)(n-m+2)C_{(2)} + \tfrac{1}{128}n(n-5)(n-m)(n-m-1)C_{(1^2)}$$

$$= \tfrac{1}{128}(n-3)(n-5)(n-m)(n-m+2)a_1^{*2} - \tfrac{1}{32}(n-5)(n-m)(m-2)a_2^{*}. \qquad \text{(A.1.14)}$$

Now a_1^{*2} can be found by squaring (2.44) and the second elementary symmetric function of the matrix $R^{-1} = (\rho_{ij})$ can be found by taking the sum of the 2×2 principal minors i.e.

$$\rho_{ij} = \sum_{u=1}^{m} \beta_i h_{ui} \alpha_u k_{ju}$$

$$a_2^{*} = \sum_{i<j} (\rho_{ii}\rho_{jj} - \rho_{ij}\rho_{ji})$$

$$= \sum_{i<j} \sum_{u,v=1}^{m} \alpha_u \alpha_v \beta_i \beta_j h_{ui} h_{vj} (k_{iu}k_{jv} - k_{ju}k_{iv}). \qquad \text{(A.1.15)}$$

On substitution for the h_{ij}, k_{ij} in terms of s_{ij}, t_{ij} if all four indices i,j,u,v are unequal then each term is clearly $O(s^4)$ and will be disregarded.

Also i=j is impossible and u=v makes the term (...) of (A.1.15) zero. Only six combinations can possibly lead to a contribution. They are

1. i=u j=v 3. i=u 5. j=u

2. i=v j=u 4. i=v 6. j=v .

Substitution in (2.40) and combination with the results (A.1.13) gives

$$G(A,B) = 1 - \tfrac{1}{6}(m-2)\sum_{i\neq j}\frac{a_i b_i}{c_{ij}} - \tfrac{1}{8}(n-3)(n-m)\sum\frac{1}{a_i b_i}$$
$$+ \tfrac{1}{48}(n-3)(n-m)(m-2)\left\{\sum_{i\neq j}\frac{a_i b_i + a_j b_j}{a_i b_i c_{ij}} + \sum_{i\neq j\neq k}\frac{a_i\ b_i}{a_k b_k c_{ij}}\right\}$$
$$+ \tfrac{1}{128}(n-3)(n-5)(n-m)(n-m+2)\sum\frac{1}{a_i{}^2 b_i{}^2}$$
$$+ \tfrac{1}{128}(n-m)(n^3-n^2m+2n^2+8m^2-10nm+23n-5m-2)\sum_{i\neq j}\frac{1}{a_i b_i a_j b_j}+\ldots \qquad \text{(A.1.16)}$$

APPENDIX 2

RAO [30] considers the case of the columns of M defining a k dimensional plane rather than a k dimensional subspace. He also assumes repeated sampling on each of n populations. The derivation of (3.23) that follows is a simplified version of the proof of result (8c.6.4) p 475.

First we need a LEMMA from [30]. Let A $m\times m$ symmetric have latent roots $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$ and corresponding latent vectors $p_1, \ldots p_m$.

LEMMA A.2.1 (1f.2.8)

Let $x_1, \ldots x_k$ be mutually orthonormal $m\times 1$ vectors. Then

$$\sup_{x_1,\ldots x_k} \sum_{i=1}^{k} x_i' A x_i = \sum_{i=1}^{k} \lambda_i$$

and the supremum is attained when $x_i = p_i$, $i=1,\ldots k$.

Now the likelihood function is, apart from a constant,

$$L(M) = \operatorname{etr}[-\tfrac{1}{2}\Sigma^{-1}(X-M)(X-M)']$$

and the likelihood ratio is

$$\lambda = \frac{\sup_{H_0} L(M)}{\sup_{H_1} L(M)} .$$

Asymptotically

$$\chi^2 = -2 \log \lambda = -2\left[\sup_{H_0} \ln L - \sup_{H_1} \ln L\right]$$

and on H_1, $\hat{M} = X$ giving $\ln L = 0$. Thus we must find

$$\chi^2 = \inf_{H_0} \text{tr}\, \Sigma^{-1}(X-M)(X-M)' . \qquad (A.2.1)$$

Make the substitutions $Y = \Sigma^{-\frac{1}{2}}X$, $N = \Sigma^{-\frac{1}{2}}M$ to give

$$\chi^2 = \inf_{H_0} \text{tr}(Y-N)(Y-N)' \qquad (A.2.2)$$

where N has rank k on H_0. Let $Y = (y_1, \ldots y_n)$, $N = (\eta_1, \ldots \eta_n)$ then (A.2.2) becomes

$$\chi^2 = \inf_{H_0} \sum_{i=1}^{n} (y_i - \eta_i)'(y_i - \eta_i) . \qquad (A.2.3)$$

If $\alpha_1, \ldots \alpha_k$ form an orthonormal basis for the space spanned by the columns of N, then

$$\eta_i = \sum_{j=1}^{k} \beta_{ij}\alpha_j = A\, \beta_i$$

where $A = (\alpha_1 \ldots \alpha_k)$ and $\beta_i' = (\beta_{i1} \ldots \beta_{ik})$. The ith term of the sum to be minimised in (A.2.3) has the form $(y_i - A\,\beta_i)'(y_i - A\,\beta_i)$ and for fixed $\alpha_1, \ldots \alpha_k$, this corresponds to the sum of squares to be minimised on the linear model $E[y_i] = A\,\beta_i$. By the usual theory, the residual sum of squares is (noting that $A'A = I_k$) $y_i'y_i - y_i'AA'y_i$. Thus we have to find

$$\min_{\alpha_1 \ldots \alpha_k} \left\{ \sum_{i=1}^{n} y_i'y_i - \sum_{i=1}^{n} y_i'AA'y_i \right\}$$

and since the first sum is a constant this reduces to finding

$$\max_{\alpha_1 \ldots \alpha_k} \sum_{i=1}^{n} y_i'AA'y_i .$$

Now

$$\sum_{i=1}^{n}(A'y_i)'(A'y_i) = \sum_{i=1}^{n}\sum_{j=1}^{k}(a_j'y_i)'(\alpha_j'y_i)$$

$$= \sum_{j=1}^{k}\alpha_j'(\sum_{i=1}^{n}y_iy_i')\alpha_j$$

and by LEMMA A.2.1

$$\max_{\alpha_1\ldots\alpha_k}\sum_{j=1}^{k}\alpha_j'(YY')\alpha_j = w_1+\ldots+w_k \qquad (A.2.4)$$

where the w_i are the latent roots of

$$\det(YY' - wI) = \det(\Sigma^{-\frac{1}{2}}XX'\Sigma^{-\frac{1}{2}} - wI) = 0.$$

Also

$$\sum_{i=1}^{n}y_i'y_i = \sum_{i=1}^{n}\text{tr } y_iy_i' = \text{tr } YY' = w_1+\ldots+w_m. \qquad (A.2.5)$$

Combining (A.2.4) and (A.2.5) we have the result

THEOREM A.2.1

$$\chi^2 = \min_{H_0} \text{tr } \Sigma^{-1}(X-M)(X-M)' = w_{k+1} +\ldots+ w_m \qquad (A.2.6)$$

where the w_i are the latent roots of $\det(XX'-w\Sigma) = 0$.

APPENDIX 3

CALCULATION OF THE COEFFICIENTS $c_{\kappa\tau}$

A.3.1 The programme

A FORTRAN computer program was written to perform the calculations outlined in section 7.2. The purpose was to calculate the coefficients $c_{\kappa\tau}$ for $k \leqslant 13$. These coefficients were written on to magnetic tape and used as input to computer programmes for summing zonal series and series of Laguerre polynomials.

A listing of the programme is given at the end of this Appendix and in the following sections important mathematical and practical features are discussed. Other features are explained by comments in the listing and by reference to the appropriate formulae.

A.3.2 The generation of partitions

It is preferable that the partitions of a given k be generated in decreasing order (section 1.4). The following algorithm is such that when given a partition it generates the one immediately below it. It is perhaps easiest understood in terms of a verbal flow chart.

Let $\kappa = (k_1 \ldots k_r)$ and its successor is $\tau = (\ell_1 \ldots \ell_s)$. Both are partitions of k and all parts are non-zero. The algorithm is initialised by presetting the first partition

$$k_1 = k, \quad r = 1.$$

1. Input the current partition $(k_1, \ldots k_r)$.
2. If $k_1 = 1$, all partitions of k have been

generated. Stop. Otherwise go to step 3.

3. If $k_r > 1$, set $s = r+1$, $\ell_u = k_u$, $u = 1,\ldots r-1$, $\ell_r = k_r - 1$, $\ell_{r+1} = 1$. This is now the new partition ready for use and storage. Afterwards return to step 1 with τ as the current partition.

4. Otherwise, find i such that $k_i > 1$, $k_{i+1} = 1$.

5. Set $\ell_u = k_u$, $u = 1,\ldots i-1$, $\ell_i = k_i - 1$.

6. The sum of the remaining k_u, $u = i+1,\ldots r$ is $r-i$ so find s,α such that

$$0 < \alpha = r-i+1-\ell_i(s-i-1) \leqslant \ell_i.$$

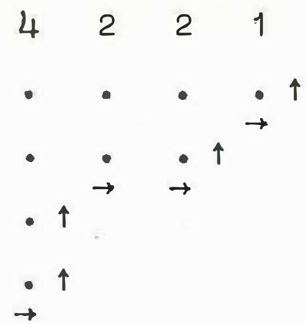
7. Set $\ell_u = \ell_i$, $u = i+1,\ldots s-1$, $\ell_s = \alpha$. Use and store this new partition and return to step 1 with τ as the current partition.

Relating this to the programme listing. The subroutine KAPPA generates the successor of the supplied partition while PSET initialises the list by setting $k_1 = k$, $r = 1$.

A.3.3 Storage of partitions

A binary representation of the partition is generated. This minimises the storage needed to record them for later use in the program.

Consider the "Young" diagram for the partition 42^21.



For convenience the diagram is written vertically instead of its usual horizontal configuration. Coding 1 for a shift right and 0 for a vertical shift, the partition 42^21 can be uniquely represented by the binary sequence 10011010. This binary sequence can then be stored in one computer word rather than using one word for each part of the partition. Incidentally the number of 1's equals the number of parts in the partition and the number of 0's is equal to the value of the largest part.

Another advantage of this binary representation is in improving search efficiency. The comparison of the actual partitions is not very convenient on a computer. This is replaced by a search of the list of binary representations (of a given k or set of k's) to see if a given binary number is on it. Thus after the partition μ is generated from τ and the elements sorted into decreasing order, the next step is to find its binary representation. It is then a simple matter to see if $\mu \leqslant \kappa$ by comparing its binary representation with the list of binary representations of all partitions of k from κ to τ.

A call to the function IBIN generates the binary representation of the partition. The function is coded in COMPASS (the assembly language for the CDC6400) a more convenient language for this type of operation. Since all parts of a partition are non-zero the binary representation ends with at least one 0. The actual representation generated by IBIN has this final 0 eliminated. A reason for doing this is given in Appendix 4.

A.3.4 The calculation of $c(\kappa)$ and $\chi_{[2\kappa]}(1)$

The normalising factor $c(\kappa)$, for converting Z_κ to C_κ, and the character $\chi_{[2\kappa]}(1)$ are calculated conveniently using formulae derived from JAMES [22] equation (3.2).

Let $\kappa = (k_1, \ldots k_r)$ be a partition of k into r non-zero parts. Then from JAMES [21] equation (21)

$$c(\kappa) = 2^k k! \frac{\prod_{i<j}^{r} (2k_i - 2k_j - i + j)}{\prod_{i=1}^{r} (2k_i + r - i)!}$$

$$\underset{(3.2)}{=} 2^k \prod_{i=1}^{r} \left(\prod_{j=1}^{k_i} \frac{A(i)}{(k_j+j)j} \prod_{j=1}^{r-i} \left(1 - \frac{2k_{i+j}}{2k_i+j}\right)\right) \quad (A.3.1)$$

where $A(i) = k_1 + \ldots + k_{i-1}$, $A(1) = 0$ and

$$\chi_{[2\kappa]}(1) = (2k)! \frac{\prod_{i<j}^{r} (2k_i - 2k_j - i + j)}{\prod_{i=1}^{r} (2k_i + r - i)!}$$

$$= \prod_{i=1}^{r} \left(\prod_{j=1}^{k_i} \frac{(B(i)+2j-1)(B(i)+2j)}{(k_i+j)j} \prod_{j=1}^{r-i} \left(1 - \frac{2k_{i+j}}{2k_i+j}\right)\right)$$

(A.3.2)

where $B(i) = 2(k_1+\ldots+k_{i-1})$, $B(1) = 0$.

The subroutine ZONCHAR is used to generate the coefficients $c(\kappa)$ using (A.3.1), and these are then stored on magnetic tape along with the $c_{\kappa\tau}$. They are used in calculating the $C_\kappa(S)$ from the $Z_\kappa(S)$ when evaluating the generalised hypergeometric functions.

The formula (A.3.2) was used to write a separate computer programme to calculate the characters listed in TABLES 7.1 and 7.2.

```
THIS PROGRAM EVALUATES THE COEFFICIENTS C(KAPPA,NU) THAT ARISE WHEN
THE ZONAL POLYNOMIALS Z(KAPPA) ARE EXPRESSED IN TERMS OF THE MONOMIAL
SYMMETRIC FUNCTIONS M(NU).
THE PROGRAM ALSO GENERATES AND STORES THE PARTITIONS
KAPPA = (K1,K2,...,KR),THEIR BINARY REPRESENTATIONS AND THE CONVERSION
FACTOR FROM Z(KAPPA) TO C(KAPPA)

      PROGRAM NY(INPUT,OUTPUT,TAPE10)
      COMMON KP(500,15),LGTH(500),CK(500),RK(500),IPOS,K,NPART
      COMMON/A/KKK(20)
      DIMENSION KQ(15),KC(15),KBS(15),CO(500)
      DIMENSION ZONCH(500),NN(500)
      DO1001I=1,20
 1001 KKK(I)=2**(I-1)
      PRINT41
   41 FORMAT(1H1)
      M=20

OUTPUT OF DATA FOR K=1

      NPART=1
      NN(1)=1
      ZONCH(1)=1
      LL=1
      KP(1,1)=1
      WRITE(10,144)NPART
      WRITE(10,142)NN(1),ZONCH(1),LL,KP(1,1)
      CO=1.
      WRITE(10,149)CO(1)
      DO80K=2,13

 GENERATE AN ORDERED LIST OF THE PARTITIONS OF K

      IPOS=0
      CALL PSET(K,KQ,M,MS,LTH)
      GOTO1
    3 CALL KAPPA(K,KQ,M,MS,LTH)
    1 IPOS=IPOS+1
      DO10I=1,LTH
   10 KP(IPOS,I)=KQ(I)
      KP(IPOS,LTH+1)=0
      LGTH(IPOS)=LTH

 CALCULATE AND STORE  C(KAPPA,KAPPA) , RHO(KAPPA) ,
 NORMALISING FACTOR AND BINARY REPRESENTATION FOR
 EACH PARTITION KAPPA

      CALL EAD
      CALL RHO
      CALL ZONCHAR(LTH,K,KQ,ZONCHR)
      ZONCH(IPOS)=ZONCHR
      NN(IPOS)=IBIN(KQ,LTH)
      IF(MS-1)3,2,3
    2 NPART=IPOS

 WRITE THE DATA ON MAGNETIC TAPE

      WRITE(10,144)NPART
  144 FORMAT(I6)
      DO20I=1,NPART
      LL=LGTH(I)
   20 WRITE(10,142)NN(I),ZONCH(I),LL,(KP(I,J),J=1,LL)
  142 FORMAT(O6,E20.13,14I3)

 STEP THROUGH THE PARTITIONS KAPPA ONE BY ONE

      NPAR=NPART-1
      DO70IZ=1,NPAR
      CO(IZ)=CK(IZ)
      LMT=LGTH(IZ)
      DO71J=1,LMT
   71 KBS(J)=KP(IZ,J)

 FOR EACH KAPPA STEP THROUGH ALL TAU = (L1,L2,...,LS)
 BELOW KAPPA

      IZZ=IZ+1
      DO30I=IZZ,NPART
      N=0
      SUM=0.
      ML=LGTH(I)
```

```
 FOR EACH TAU GENERATE ALL POSSIBLE

      MU = (L1...LI+T...LJ-T...LS)

   MM=ML-1
   DO50IA=2,ML
   ID=IA-1
   DO50IB=1,ID
   MN=KP(I,IA)
   DO60IC=1,MN
   DO40J=1,ML
40 KC(J)=KP(I,J)
   KC(IB)=KC(IB)+IC
   KC(IA)=KC(IA)-IC
   KK=KC(IB)-KC(IA)

 FOR EACH MU SORT THE ELEMENTS INTO DECREASING ORDER
 AND FIND THE BINARY REPRESENTATION

   CALL SHORT(KC,ML)
   IF(KC(ML).EQ.0)GOTO12
   MLT=ML
   GOTO18
12 MLT=ML-1
18 NNN=IBIN(KC,MLT)

 TEST IF THE SORTED MU IS ADMISSIBLE
 IF TRUE ADD APPROPRIATE TERM TO THE SUM AS PER (7.6)

   IS=I-1
   DO73J=IZ,IS
   IF(NNN-NN(J))73,38,73
73 CONTINUE
   GOTO60
38 SUM=SUM+KK*CO(J)
   N=N+1
60 CONTINUE
50 CONTINUE
   IF(N)81,82,81
82 CO(I)=0.
   GOTO30
81 IF(ABS(SUM)-0.5)82,82,83
83 CO(I)=SUM/(RK(IZ)-RK(I))
30 CONTINUE

 OUTPUT THE COEFFICIENTS ONTO MAGNETIC TAPE

    WRITE(10,149)(CO(LL),LL=IZ,NPART)
149 FORMAT(9F15.1)
 70 CONTINUE
    WRITE(10,149)CK(NPART)
 80 CONTINUE
    STOP $ END


 CALCULATES THE LEADING COEFFICIENT C(KAPPA,KAPPA) ACCORDING
 TO (7.5)

   SUBROUTINE EAD
   COMMON KP(500,15),LGTH(500),CK(500),RK(500),IPOS,K,NPART
   PR=1
   LTH=LGTH(IPOS)
   DO10I=1,LTH
   KL=KP(IPOS,I)
   LL=KL-KP(IPOS,I+1)
   DO10J=1,I
   A=(I-J)/2.+(KP(IPOS,J)-KL)+0.5
10 PR=PR*COF(A,LL)
   CK(IPOS)=PR*2.**K
   RETURN $ END


 CALCULATES THE (A)  = A(A+1)...(A+K-1)
                   K

   FUNCTION COF(A,K)
   C=A
   II=1
 3 IF(II-K)1,2,4
 1 C=C*(A+II)
   II=II+1
   GOTO3
 4 COF=1.
   RETURN
 2 COF=C
   RETURN $ END
```

```
   CALCULATES THE RHO(KAPPA) ACCORDING TO (7.7)

      SUBROUTINE PHO
      COMMON KP(500,15),LGTH(500),CK(500),RK(500),IPOS,K,NPART
      SUM=0
      I=LGTH(IPOS)
      DO10J=1,I
      A=KP(IPOS,J)
   10 SUM=SUM+A*(A-J)
      RK(IPOS)=SUM
      RETURN $ END


   CALCULATES THE NORMALISING FACTOR FOR CONVERSION FROM
   Z(KAPPA) TO C(KAPPA) USING FORMULA (A.3.1)

      SUBROUTINE ZONCHAR(P,K,KAPPA,ZONCHR)
      INTEGER P,K,KAPPA,KAPK,Q
      DIMENSION KAPPA(P)
      REAL ZONCHR,PR
      KAPK=0.
      PR=1.
      DO10I=1,P
      KAPPAI=KAPPA(I)
      DO20J=1,KAPPAI
      PR=PR*FLOAT(KAPK+J)/(FLOAT(KAPPAI+J)*FLOAT(J))
   20 CONTINUE
      IF(I.EQ.P)1,2
    2 Q=P-I
      DO30J=1,Q
      PR=PR*(1.-FLOAT(2*KAPPA(I+J))/FLOAT(2*KAPPAI+J))
   30 CONTINUE
    1 KAPK=KAPK+KAPPAI
   10 CONTINUE
      ZONCHR=2**K*PR
      RETURN $ END


      SUBROUTINE PSET(K,KP,M,MS,LGTH)
      DIMENSION KP(1)
      KP = K
      DO 10 I=2,K
10    KP(I) = 0
      MS = 0
      LGTH = 1
      RETURN
      ENTRY KAPPA
      IF(K.EQ.1) GO TO 37
      MS = 0
9     N = 0
      DO 11 KM = 1,K
      J = K-KM+1
      N = N + KP(J)
      IF(KP(J).GT.1) GO TO 36
11    CONTINUE
   37 CONTINUE
      MS = 1
      LGTH = K
      RETURN
36    CONTINUE
3     IF(KP(J+1).GT.0) GO TO 4
      KP(J) = KP(J) - 1
      KP(J+1) = 1
      LGTH = J+1
      IF(LGTH.GT.M) GO TO 9
      IF(KP.EQ.1) GO TO 37
      RETURN
4     KP(J) = LD = KP(J)-1
      MM = J
6     MM = MM + 1
      N = N-LD
      IF(N.LT.0) GO TO 5
      KP(MM) = LD
      GO TO 6
5     N = N+LD
      MM = MM - 1
      KP(MM) = N
      LGTH = MM
      IF(N.EQ.0) LGTH = LGTH - 1
      MM = MM + 1
      DO 27 IM = MM,K
   27 KP(IM) = 0
      IF(LGTH.GT.M) GO TO 9
      IF(KP.EQ.1) GO TO 37
      RETURN
      END
```

CALCULATES THE BINARY REPRESENTATION OF A PARTITION

```
         IDENT    IBIN
         ENTRY    IBIN
         USE      /A/
KK       BSS      20
         USE      *
IBIN     BSS      1
         SB4      1
         SA1      B2
         SB2      X1+B1
         SA2      B2-B4
         SA1      X2+KK-1
ST       SB2      B2-B4
         EQ       B1,B2,RF
         SA3      B2-B4
         IX2      X2+X3
         SA3      B2
         IX2      X2-X3
         SX2      X2+B4
         SA3      X2+KK-1
         IX1      X1+X3
         JP       ST
RF       BX6      X1
         JP       IBIN
```

CALCULATES THE PARTITION FROM ITS BINARY REPRESENTATION

```
         ENTRY    NIBI
NIBI     BSS      1
         SB4      1
         SA1      B1
         SB1      X1
         SA2      B2
         SX6      B4
         SX3      B4
LL       ZR       B1,EN
         SB1      B1-B4
L        BX4      X2*X3
         ZR       X4,JU
         SA6      B3+B1
         LX3      1
         JP       LL
JU       SX6      X6+B4
         LX3      1
         JP       L
EN       SA6      B3
         JP       NIBI
         END
```

SORTS THE ELEMENTS L1...LI+T...LJ-T...LS OF MU INTO DECREASING ORDER

```
         IDENT    SHORT
         ENTRY    SHORT
SHORT    BSS      1
         SA1      B2
         SB2      X1-1
EL       SB3      B0
         SB4      B0
ES       EQ       B2,B3,ER
         SA2      B1+B3
         SA3      A2+1
         IX4      X2-X3
         PL       X4,EK
         BX6      X2
         BX7      X3
         SA6      A3
         SA7      A2
         SB4      B4+1
EK       SB3      B3+1
         JP       ES
ER       ZR       B4,SHORT
         SB2      B2-1
         ZR       B2,SHORT
         JP       EL
         END
```

APPENDIX 4

CALCULATION OF THE $g^{\kappa}_{\nu\mu}$ AND $\binom{\kappa}{\nu}$

A.4.1 The programme

The FORTRAN programme listed at the end of this Appendix is designed to generate the coefficients $g^{\kappa}_{\nu\mu}$ and $\binom{\kappa}{\nu}$ for $k \leqslant 9$ using the formulae of sections 7.4 and 7.5. Only the coefficients $\binom{\kappa}{\nu}$ are retained on magnetic tape. These are used in the calculation of the Laguerre polynomials. The coefficients $g^{\kappa}_{\nu\mu}$ are not saved as they are only used here to calculate the $\binom{\kappa}{\nu}$ and as indicated in section 7.5 they have already been tabled for $k \leqslant 7$. Basically the programme is designed to calculate the $g^{\kappa}_{\nu\mu}$ and it is then a simple matter to selectively sum them to derive the $\binom{\kappa}{\nu}$.

The evaluation of the b_ε and an efficient method of storing them are discussed in the next two sections.

A.4.2 The product of msf's, the b_ε and the $e^{\delta}_{\alpha\beta}$

Let α be a partition of r with m parts
β be a partition of t with n parts
ε be a partition of k with p parts

under the conditions $k = r+s$, $m \geqslant n$, $m \leqslant p \leqslant m+n$. Let $\alpha = (f_1,\ldots f_m)$, $\beta = (h_1,\ldots h_n)$, $\varepsilon = (\ell_1,\ldots \ell_p)$. It is important that all parts be non-zero. Zero parts would lead to unwanted extra terms.

As for (7.34) and (7.35) the product of msf's is

$$M_\alpha M_\beta = \{\Sigma s_i^{f_1} s_j^{f_2} \ldots s_q^{f_m}\}\{\Sigma s_u^{h_1} s_v^{h_2} \ldots s_w^{h_n}\} \quad (A.4.1)$$

$$= \sum_{\varepsilon \leqslant \tau} b_\varepsilon M_\varepsilon \quad (A.4.2)$$

where $\tau = (f_1+h_1, f_2+h_2, \ldots f_n+h_n, f_{n+1}, \ldots f_m)$. The msf's M_ε are formed by taking all possible products in (A.4.1).

At the i^{th} stage:

Select i elements from $f_1, \ldots f_m$ to give the subset $S_i(\alpha)$.
Select i elements from $h_1, \ldots h_n$ to give the subset $S_i(\beta)$.
Arrange the elements of $S_i(\alpha)$ in some order and hold it fixed. Then permute the elements of $S_i(\beta)$ in all possible ways and after each permutation add these elements to the corresponding ones of $S_i(\alpha)$. To each generated list append the $m-i$ remaining elements of α and $n-i$ of β to form a partition ε. Each partition ε is associated with a possible msf M_ε from (A.4.1).

For example $m = 5$, $n = 4$, $i = 3$:

$S_3(\alpha) = f_1, f_3, f_5$

$S_3(\beta) = h_1, h_2, h_4$.

Two possible permutations of $S_3(\beta)$ are

h_1, h_4, h_2 and h_2, h_1, h_4

to give

$$\varepsilon_1 = f_1+h_1, f_3+h_4, f_5+h_2, f_2, f_4, h_3$$

$$\varepsilon_2 = f_1+h_2, f_3+h_1, f_5+h_4, f_2, f_4, h_3.$$

Summarising, for fixed α and β we may have $i = 0, 1, \ldots n$. For each i we generate all i element subsets $S_i(\alpha)$ and for each $S_i(\alpha)$ we generate all possible i element subsets $S_i(\beta)$. For each pair of $S_i(\alpha)$ and $S_i(\beta)$ we generate all possible permutations of $S_i(\beta)$ before adding it to $S_i(\alpha)$. The remaining m+n-2i elements of α and β are appended to generate an ε. Using (7.36) the b_ε is calculated.

After all acceptable combinations of $S_i(\alpha)$ and $S_i(\beta)$ are used we are left with the $e^{\delta}_{\alpha\beta}$ for the product of M_α and M_β.

The selection of all possible subsets $S_i(\alpha)$ for a partition α is done by the pair of subroutines SELSET and SELGET. The routine SELSET initialises by setting $S_i(\alpha) = f_1,\ldots f_i$ while the routine SELGET generates a new $S_i(\alpha)$ on each call. Similarly for all $S_i(\beta)$ of β. An example suffices to illustrate the principle.

Set m = 5, i = 3.

SELSET		f_1, f_2, f_3
SELGET	1.	f_1, f_2, f_4
	2.	f_1, f_2, f_5
	3.	f_1, f_3, f_4
	4.	f_1, f_3, f_5
	etc....	

The permutation of $S_i(\beta)$ is performed by the pair of subroutines PERSET,PERGET. Subroutine PERSET initialises arrays and returns the identity permutation, while the routine PERGET generates a new permutation from the current one using an algorithm given by LEHMER (p 23) in BECKENBACH [6].

A.4.3 Storing the b_ε

Every time a new b_ε is generated this must be added to the accumulated total associated with that partition. One method of storage is to convert ε to its binary representation and use this to search a list of the binary representations of all partitions κ of k and then increment the associated coefficient. This is slow and inefficient as much searching is involved.

A much faster method is to use the binary representation of ε as an index. For example:

partition	binary representation	index (base 10)
4	1000	8
31	1001	9
2^2	110	6
21^2	1011	11
1^4	1111	15

(the final 0 has been dropped as per Appendix 3).

Thus any $b_{(4)}$ is added to storage location 8

any $b_{(31)}$ is added to storage location 9

etc... .

As stated before each κ has a unique index. The largest index for $\kappa \leqslant 9$ is 511 (corresponding to 1^9 having the binary representation 111111111). Only some of the storage locations are used for $k \leqslant 9$, but the increase in speed of operation far outweighs the disadvantage of needing 511 storage locations reserved for the 97 partitions of $k = 1, \ldots 9$.

```
      THIS PROGRAM CALCULATES THE ZONAL POLYNOMIAL PRODUCT AND
      GENERALISED BINOMIAL COEFFICIENTS ALONG THE LINES OF
      SECTIONS 7.4 AND 7.5
      THE GENERALISED BINOMIAL COEFFICIENTS ARE STORED ON
      MAGNETIC TAPE FOR FUTURE USE

        PROGRAM AKL(INPUT.OUTPUT.TAPE10.TAPE20)
        COMMON NN(140).LTH(140).COR(1860).CO(4032).KA(10).KB(10).KC(10).
       *Z(42)
        COMMON/STAR/KCC(10)
        COMMON/START/NP(11).NB(42).NBS(11)
        DATA NB/0.1.3.6.10.15.21.28.36.45.55.66.78.91.105.120.136.153.
       *171.190.210.231.253.276.300.325.351.378.406.435.465.496.528.561.
       *595.630.666.703.741.780.820.861/
        DATA NP/1.2.4.7.12.19.30.45.67.97.139/
        DATA NBS/0.1.4.10.25.53.119.239.492.957.1860/
        DATA KCC/0.2.4.30.77.198.435.968.1980.4032/
        DIMENSION NPAIR(142.2).N(202)
        COMMON APR(8484)
        DIMENSION LX(10).LY(10)
        COMMON/FA/FAC(10)
        COMMON/A/KK(20)
        DO1001I=1.20
  1001  KK(I)=2**(I-1)
        FAC(1)=1. $ FAC(2)=2. $ FAC(3)=6. $ FAC(4)=24. $ FAC(5)=120.
        FAC(6)=720. $ FAC(7)=5040. $ FAC(8)=40320. $ FAC(9)=362880.
        FAC(10)=3628800.
   149  FORMAT(9F15.1)
   144  FORMAT(I6)
   142  FORMAT(O6.F20.13.I3)

      INPUT THE DATA WRITTEN ON MAGNETIC TAPE BY THE PROGRAM
      DESCRIBED IN APPENDIX 3
      CALCULATE THE COEFFICIENTS A(KAPPA.TAU) OF (7.27)

        IPOS=0
        DO10IJK=1.10
        READ(10.144)NPART
        DO20J=1.NPART
        IPOS=IPOS+1
    20  READ(10.142)NN(IPOS).Z(J).LTH(IPOS)
        II=NBS(IJK)
        DO30I=1.NPART
        READ(10.149)(CO(LL).LL=1.NPART)
        DO40J=1.NPART
        LL=II+I+NB(J)
    40  COR(LL)=CO(J)*Z(I)
    30  CONTINUE
    10  CONTINUE
        N(1)=0
        DO99I=1.201
    99  N(I+1)=N(I)+42
        PRINT41
    41  FORMAT(1H1)
        DO50I=2.10
        IPR=0
        DO51IXYZ=1.8484
    51  APR(IXYZ)=0.
        IX=NP(I)
        IY=NP(I+1)-1
        J=I
     4  J=J-1
        K=I-J
        IXX=0
        IF(J-K)52.2.3
     2  IXX=1
     3  IA=NP(J+1)-NP(J)
        IB=NP(K+1)-NP(K)
        IXX=IXX+1
        DO60JA=1.IA
        GOTO(61.62).IXX
    61  IYY=1
        GOTO63
    62  IYY=JA
    63  DO60JB=IYY.IB
        II=NBS(J)
        LLL=NBS(K)
        MM=NBS(I)
        JS=NP(J)+JA-1
        JST=NP(J+1)-1
        KS=NP(K)+JB-1
        KST=NP(K+1)-1
        IPR=IPR+1
        NPAIR(IPR.1)=NN(JS)
        NPAIR(IPR.2)=NN(KS)
        DO70MA=JS.JST
        DO70MB=KS.KST
```

```
C     SELECT A PAIR OF PARTITIONS
C
C          ALPHA  A PARTITION OF R
C          BETA   A PARTITION OF 1
C
      LA=LTH(MA)
      LB=LTH(MB)
      IF(LA-LB)68,69,69
   68 CALL NIBI(LA,NN(MA),KB)
      CALL NIBI(LB,NN(MB),KA)
      LA=LB
      LB=LTH(MA)
      GOTO67
   69 CALL NIBI(LA,NN(MA),KA)
      CALL NIBI(LB,NN(MB),KB)
   67 MMA=MA-NP(J)+1
      MMB=MB-NP(K)+1
      LLA=II+JA+NB(MMA)
      LLB=LLL+JB+NB(MMB)
      AC=COR(LLA)*COR(LLB)
      FACA=FRAC(KA,LA)*FRAC(KB,LB)
      FACB=AC/FACA
      DO110IL=1,LA
  110 KC(IL)=KA(IL)
      DO120IL=1,LB
  120 KC(LA+IL)=KB(IL)
      LAB=LA+LB
      CALL SHORT(KC,LAB)
      NNN=IBIN(KC,LAB)
      LL=INDEX(NNN,1)+N(IPR)
      APR(LL)=FRAC(KC,LAB)*FACB+APR(LL)
      DO130JJ=1,LB
      CALL SELSET(LX,JJ,M,MS,LA)
      GOTO131
  134 CALL SELGET(LX,JJ,M,MS,LA)
      IF(MS)130,131,130
  131 CALL SELSET(LY,JJ,MY,MSY,LB)
      GOTO133
  138 CALL SELGET(LY,JJ,MY,MSY,LB)
      IF(MSY)134,133,134
  133 IF(JJ-1)136,135,136
  135 DO140IL=1,LA
  140 KC(IL)=KA(IL)
      JJA=LX(1)
      JJB=LY(1)
      KC(JJA)=KC(JJA)+KB(JJB)
      JJC=LA
      DO150IL=1,LB
      IF(IL-JJB)151,150,151
  151 JJC=JJC+1
      KC(JJC)=KB(IL)
  150 CONTINUE
      CALL SHORT(KC,JJC)
      NNN=IBIN(KC,JJC)
      LL=INDEX(NNN,1)+N(IPR)
      APR(LL)=FRAC(KC,JJC)*FACB+APR(LL)
      GOTO138
  136 CALL PERSET(LY,JJ,MSY)
      GOTO137
  139 CALL PERGET(LY,JJ,MSX)
      IF(MSX)138,137,138
  137 DO160IL=1,LA
  160 KC(IL)=KA(IL)
      DO170IL=1,JJ
      JJA=LX(IL)
      JJB=LY(IL)
  170 KC(JJA)=KC(JJA)+KB(JJB)
      JJC=LA
      DO180IL=1,LB
      DO190ILL=1,JJ
      IF(IL-LY(ILL))190,181,190
  190 CONTINUE
      JJC=JJC+1
      KC(JJC)=KB(IL)
  181 CONTINUE
  180 CONTINUE
      CALL SHORT(KC,JJC)
      NNN=IBIN(KC,JJC)
      LL=INDEX(NNN,1)+N(IPR)
      APR(LL)=FRAC(KC,JJC)*FACB+APR(LL)
      GOTO139
  130 CONTINUE
   70 CONTINUE
      NPART=NP(I+1)-NP(I)
      IM=N(IPR)
      NPAR=NPART-1
```

```
      DO210MA=1,NPAR
      MBX=MA+1
      LL=IM+MA
      LLX=MM+MA+NB(MA)
      APR(LL)=ZR=APR(LL)/COR(LLX)
      DO220MB=MBX,NPAPT
      LL=IM+MB
      LLX=MM+MA+NB(MB)
  220 APR(LL)=APR(LL)-ZR*COR(LLX)
  210 CONTINUE
      LL=IM+NPART
      LLX=MM+NPART+NB(NPART)
      APR(LL)=APR(LL)/COR(LLX)
      IMM=IM+NPART
      IM=IM+1
      DO230LL=IM,IMM
      IF(ABS(APR(LL))-1.E-10)231,231,230
  231 APR(LL)=0.
  230 CONTINUE
   60 CONTINUE
      IF(J-K)52,52,4
   52 Z=FAC(I)
      IX=NP(I)
      IY=NP(I+1)-1
      NPART=IY-IX+1
      II=I-1
      MCAR=0
      KCCC=KCC(I)
      DO310IA=1,KCCC
  310 CO(IA)=0.
      DO320JJ=1,II
      ZZ=Z/(FAC(JJ)*FAC(I-JJ))
      IX=NP(JJ)
      IY=NP(JJ+1)-1
      DO330MA=IX,IY
      NNN=NN(MA)
      DO340MB=1,IPR
      IF(NNN-NPAIR(MB,1))341,342,341
  341 IF(NNN-NPAIR(MB,2))340,342,340
  342 DO350MC=1,NPART
      LL=N(MB)+MC
      IF(ABS(APR(LL))-1.E-10)350,350,351
  351 MCC=MCAR+MC
      CO(MCC)=CO(MCC)+APR(LL)
  350 CONTINUE
  340 CONTINUE
      DO360MC=1,NPART
      MCC=MCAR+MC
  360 CO(MCC)=CO(MCC)*ZZ
      MCAR=MCAR+NPART
  330 CONTINUE
  320 CONTINUE
      WRITE(20)(CO(LL),LL=1,KCCC)
      PRINT241,(CO(LL),LL=1,KCCC)
  241 FORMAT(X10F13.8)
   50 CONTINUE
      STOP $ END


      SUBROUTINE PERSET(L,N,MS)
      DIMENSION D(10),E(10),A(10),L(10)
      INTEGER A,D,E
      MS=0
      NN=N-1
      DO10I=1,NN
      D(I)=0
      E(I)=1
      A(I)=I+1
   10 CONTINUE
    7 RETURN
      ENTRY PERGET
      J=NN
    8 A(J)=A(J)-E(J)
      IF(A(J)-J-1)1,2,1
    1 IF(A(J))3,2,3
    3 I=J+1
      K=A(J)
    6 IF(I-NN)4,4,5
    4 K=K+D(I)
      I=I+1
      GOTO6
    5 IH=L(K)
      L(K)=L(K+1)
      L(K+1)=IH
      GOTO7
    2 E(J)=-E(J)
      D(J)=1-D(J)
      J=J-1
      IF(J)9,9,8
    9 MS=1
      IH=L(2)
      L(2)=L(1)
      L(1)=IH
      RETURN $ END
```

```
      SUBROUTINE SELSET(L,I,M,MS,N)
      DIMENSION L(1)
      MS=0
      L(I+1)=N+1
      M=1
      L(1)=1
    3 IF(M-I)1,2,1
    1 M=M+1
      L(M)=L(M-1)+1
      GOTO3
    2 RETURN
      ENTRY SELGET
    6 IF(L(M)+1-L(M+1))5,4,5
    5 L(M)=L(M)+1
      GOTO3
    4 M=M-1
      IF(M)6,7,6
    7 MS=1
      RETURN $ END
```

CHECKS THE LIST OF BINARY REPRESENTATIONS OF PARTITIONS
KAPPA AND RETURNS THE POSITION OF EPSILON ON IT

```
          IDENT     INDEX
          ENTRY     INDEX
          USE       //
NN        BSS       140
          USE       *
          USE       /START/
NP        BSS       11
NB        BSS       53
          USE       *
INDEX     BSS       1
          SA1       B1
          SA2       B2
          SA3       X2+NP-1
          SB2       X3+NN-2
          SB3       1
AG        SA2       B2+B3
          IX3       X1-X2
          ZR        X3,RM
          SB3       B3+1
          JP        AG
RM        SX6       B3
          JP        INDEX
          END
```

CALCULATES FAC(PHI1)*FAC(PHI2)*... AS REQUIRED IN (7.36)

```
          IDENT     FRAC
          ENTRY     FRAC
          USE       /FA/
FAC       BSS       10
          USE       *
ONE       DATA      1.
FRAC      BSS       1
          SA1       ONE
          BX6       X1
          SB3       B0
          SA2       B2
          SB2       X2
FV        SB2       B2-1
          ZR        B2,FR
          SA1       B1+B2
          SA2       A1-1
          IX3       X1-X2
          ZR        X3,FL
          SA5       B3+FAC
          FX6       X5*X6
          SB3       B0
          JP        FV
FL        SB3       B3+1
          JP        FV
FR        SA5       B3+FAC
          FX6       X5*X6
          JP        FRAC
          END
```

THIS PROGRAM ALSO CALLS IBIN AND SHORT
THESE HAVE ALREADY BEEN LISTED

BIBLIOGRAPHY

1. G.A. ANDERSON: An asymptotic expansion for the distribution of the latent roots of the estimated covariance matrix, Ann. Math. Statist. 36 (1965), 1153-1173.

2. T.W. ANDERSON: The non-central Wishart distribution and certain problems of multivariate statistics, Ann. Math. Statist. 17 (1946), 409-431.

3. T.W. ANDERSON and M.A. GIRSHICK: Some extensions of the Wishart distribution, Ann. Math. Statist. 15 (1944), 345-357.

4. M.S. BARTLETT: Multivariate analysis (with discussion), J.R. Statist. Soc. (Suppl.) 9 (1947), 176-190.

5. M.S. BARTLETT: A note on the multiplying factors for various χ^2 approximations, J.R. Statist. Soc. B 16 (1954), 296-298.

6. EDWIN F. BECKENBACH: "Applied Combinatorial Mathematics", University of California Engineering and Physical Sciences Extension Series (1964), John Wiley and Sons, Inc.

7. A.G. CONSTANTINE: Some non-central distribution problems in multivariate analysis, Ann. Math. Statist. 34 (1963), 1270-1285.

8. A.G. CONSTANTINE: The distribution of Hotelling's generalised T_0^2, Ann. Math. Statist. 37 (1966), 215-225.

9. ARTHUR ERDELYI et al.: "Higher Transcendental Functions", McGraw Hill (1953), New York.

10. A. ERDELYI: "Asymptotic Expansions", Dover (1956), New York.

11. R.A. FISHER: The sampling distribution of some statistics obtained from non-linear equations", Ann. Eugenics 9 (1939), 238-249.

12. C.S. HERZ: Bessel functions of matrix argument, Ann. of Math. 65 (1955), 474-523.

13. P.L. HSU: On the distribution of the roots of certain determinantal equations, Ann. Eugenics 9 (1939), 250-258.

14. L.C. HSU: A theorem on the asymptotic behaviour of a multiple integral, Duke Math. J. 15 (1948), 623-632.

15. A.T. JAMES: Normal multivariate analysis and the orthogonal group, Ann. Math. Statist. 25 (1954), 40-75.

16. A.T. JAMES: The non-central Wishart distribution, Proc. Roy. Soc. (London) Ser. A 229 (1955), 364-366.

17. A.T. JAMES: A generating function for averages over the orthogonal group, Proc. Roy. Soc. (London) Ser. A 229 (1955), 367-375.

18. A.T. JAMES: The distribution of the latent roots of the covariance matrix, Ann. Math. Statist. 31 (1960), 151-158.

19. A.T. JAMES: Zonal polynomials of the real positive definite symmetric matrices, Ann. of Math. 74 (1961), 456-469.

20. A.T. JAMES: The distribution of non-central means with known covariance, Ann. Math. Statist. 32 (1961), 874-882.

21. A.T. JAMES: Distribution of matrix variates and latent roots derived from normal samples, Ann. Math. Statist. 35 (1964), 475-501.

22. A.T. JAMES: Inference on latent roots by calculation of hypergeometric functions of matrix argument, "**Multivariate Analysis**", Ed. by P.R. Krishnaiah, Academic Press (1966).

23. A.T. JAMES: Tests of equality of latent roots of the covariance matrix, International Symposium (July 1968), Dayton, Ohio.

24. A.T. JAMES: Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator, Ann. Math. Statist. 39 (1968), 1711-1718.

25. C.G. KHATRI and K.C.S. PILLAI: On the non-central distributions of two test criteria in multivariate analysis of variance, Ann. Math. Statist. 39 (1968), 215-226.

26. D.N. LAWLEY: Tests of significance in canonical analysis, Biometrika 46 (1959), 59-66.

27. R.J. MUIRHEAD: Systems of partial differential equations for hypergeometric functions of matrix argument, (submitted to Ann. Math. Statist.).

28. K.C.S. PILLAI and A.K. GUPTA: On the non-central distribution of the second elementary symmetric function of the roots of a matrix, Ann. Math. Statist. 39 (1968), 833-839.

29. EARL D. RAINVILLE: "Special Functions" p 212, The Macmillan Company (1960), New York.

30. C.R. RAO: "Linear Statistical Inference and its Applications", John Wiley and Sons (1965).

31. S.N. ROY: p-statistics or some generalisations in analysis of variance appropriate to multivariate problems, Sankhyā 4 (1939), 381-396.

32. B.L. van der WAERDEN: "Modern Algebra. Volume I" (English translation), Frederick Ungar Publishing Co. (1949), New York.

33. M. WEIBULL: The distribution of t- and F-statistics and of correlation and regression coefficients in stratified samples from normal populations with different means, Skand. Aktuarietidskr. 36 Supple. (1953), 1-106.

34. J. WISHART: The generalised product moment distribution in samples from a normal multivariate population, Biometrika 20 (1928), 32-43.