CIMAT
Centro de Investigación en Matemáticas, A.C.

Noncentral Elliptical Configuration Density.

TESIS
Que para obtener el grado de:
Doctor en ciencias
Con orientación en:
Matemáticas Básicas
Presenta:

Francisco José Caro Lopera.

Guanajuato, Gto. México. Agosto de 2008.

Center for Mathematical Research. CIMAT, A.C.

Noncentral Elliptical Configuration Density.

Doctoral Thesis.

Francisco J. Caro Lopera.

Department of Basic Mathematics.

CIMAT A.C.

Thesis advisors (a.o):

Dr. José A. Díaz-García.

Department of Statistics and Computation. UAAAN.

Dra. Graciela González-Farías.

Department of Probability and Statistics. CIMAT A.C.

Guanajuato, Gto., México.

August, 2008.

To Sarita, Graciela and José.

Acknowledgment

The major of the gratitudes to Dra. Graciela González Farías for all the managements and received supports to carry out my studies of doctorate.

To my thesis advisors, Dra. Graciela González Farías and Dr. José Antonio Díaz García, for the constant support received in all the aspects, academic, scientist, personal, economic; in general by their friendship and the great bulwark that they represented in my studies in Mexico. In the same lines of gratitude, for the extraordinary support of Dr. Rogelio Ramos Quiroga in these three years.

To the referees of my thesis, professors: Dr. Víctor Manuel Pérez Abreu Carrión, Dr. Óscar Adolfo Sánchez Valenzuela, Dr. Narayanaswamy Balakrishnan, Dr. Fernando Brambila Paz and Dra. Graciela María de los Dolores González Farías, for their valuable comments, revisions, corrections and important suggestions and recommendations in order to improve the thesis and potentiate the future research work in this area. Special thanks to my advisor, Dr. José Antonio Díaz García for the endless hours of work, revision, comments, suggestions about my thesis and our research papers in this three years. Also, thanks to Dr. Pedro Luis del Ángel Rodríguez for reading and comment the thesis.

Also, I want to thank the constant support of professor Stephanie Dunbar.

Thanks to CIMAT for the scholarships received during my studies and thanks to CONACYT for the doctorate scholarship (Registro de Becario

208286).

Thanks to the CONACYT project grant number P 45974- F.

Contents

Abstract

The noncentral configuration density under an elliptical model is derived; it generalizes and corrects the Gaussian configuration and some Pearson results. Then by using partition theory a number of explicit configuration densities are obtained; i.e. configuration densities associated with the matrix variate symmetric Kotz type distributions (it includes normal), the matrix variate Pearson type VII distributions (it includes t and Cauchy distributions), the matrix variate symmetric Bessel distribution (it includes Laplace distribution) and the so called matrix variate symmetric Jensen-logistic distribution.

The inference procedure for any elliptical configuration density is set in this work in terms of published efficient algorithms involving infinite confluent hypergeometric type series of zonal polynomials. The finite configuration density study is proposed and it is applied in a finite Kotz configuration density subfamily, including normal; as a consequence the inference procedure is extremely simplified because it does not require any approximation of the corresponding configuration density. Then, the applications are based on low degree zonal polynomials, computed by our formulae, and they include Biology (mouse vertebra, gorilla skulls, girl and boy craniofacial studies), Medicine (brain MR scans of schizophrenic patients) and image analysis (postcode recognition).

Preface

When the statistical theory of shape was placed in the setting of the non-central multivariate analysis (Goodall and Mardia (1993)), a wide gamma of standard theories developed in the last 60 years, were available, to solve the new distributional problems.

As usual, the first works assumed Gaussian distributions for the landmark components (Goodall and Mardia (1993), Díaz-García *et al.* (2003)), and integration over Euclidean and affine transformations provided the required shape and configuration distributions, respectively, in terms of a well study theory, the zonal polynomials of matrix argument.

Theory of integration over orthogonal and positive definite matrices involving zonal polynomials led exact distributions, but the problem for large computations remained open for years, and the use of approximations were needed for applications. Recently, with the appearance of efficient algorithms for zonal and hypergeometric functions, the exact distributions can be studied in the corresponding inference problem, and then the applications can be potentially improved (Koev and Edelman (2006)).

However, the normal constraint stands ideal, so new enriched distributions, for example, the elliptically contoured distributions, can be considered for the landmark components, but the corresponding new integrals, under the Euclidean or affine transformations, demand new developments. The Euclidean case was solved with the usual multivariate analysis and classical integration formulae for zonal and invariant polynomials of several matrix arguments (Goodall and Mardia (1993), Díaz-García *et al.* (2003), etc.). But the configuration distribution (based on affine transformations) for any elliptical model, has not been studied in literature.

Two motivations seem reasonable for solving the configuration problem, one, the geometric meaning for the applications and second, the involved distributional problem. The first one, is clearly the most important for users of shape theory; the transformation refers problems which are not equally deformed in all directions (as in the Euclidean transformation models), but uniformly deformed axe by axe. It is specially useful in growth theory, mechanical deformations, non rigid evolution, electrophoretic gel studies, etc. but even the classical applications studied by Euclidean transformations and Gaussian distributions, can be research again by guessing an elliptical model, previously ratified by a Schwarz's dimensional criterion (Schwarz (1978)), for example, and under an affine transformation. The second topic, is a mathematical problem and it considers fundamentally advances in integration over positive definite matrices via zonal polynomials. Noncentral multivariate analysis gives the key for studying one by one the particular elliptical models, by solving the corresponding multiple integral, as in the normal case (Goodall and Mardia (1993),Díaz-García *et al.* (2003)) and some integrals given for Pearson models (Xu and Fang (1989)), for example. But, this technique provides no solution for any general model and certainly, some multiple integrals seem tedious. So, an interesting approach could go in that general direction, by simplifying the integration problem.

Fortunately, some general results for integration on positive definite matrices are available (Teng *et al* (1989)) and they can be generalized and used in the setting we need for the configuration distributions. We give all the details and search back the necessary tools for showing those results from a comprehensive and self-content point of view. Clearly, those results deserve a detailed treatment because they lead that our distributional results for any elliptical model can be reduced to a computation of a single integral.

However, the simplicity in the integration will have a price, k-th derivative expressions for the elliptical model functions. Then the second tool for solving the problem appears, a partitional treatment for expressing the derivative in a fashion that the single integral can be computed easily.

With the distributional problem solved, another important question results, the computation, but as we mentioned, series of zonal polynomials can be now computed efficiently, then the inference problem based on the exact configuration density is set in this work as a solvable numerical aspect.

But we want to go further, the series for classical elliptical families, seem preserve a remarkable property of the Gaussian families, this is, the series can be finite by using some generalized Kummer relations. We just explore them for a subfamily of the Kotz distribution, which will support the applications, and they will imply that the inference problem just depends on the optimization of a low degree polynomial.

Now, in the next lines we give more details about our thesis problem. Mathematical and statistical theories of shape have been studied extensively in two decades by a number of authors; classical treatments which give excellent surveys about those topics are Kendall *et al.* (1999), Dryden and Mardia (1998), Small (1996) and the references there in.

In the real statistical approach we can find fundamentally three techniques: Shape Theory via QR Decomposition, which was founded by Goodall and Mardia (1993) and the posterior works of these authors; Shape Theory via SVD Decomposition, mainly studied by Le and Kendall (1993), Díaz-García *et al.* (2003) and subsequent works; and Shape Theory via Configuration Density, studied by the respective groups of Goodall and Díaz, and many others, see Dryden and Mardia (1998).

Goodall and Mardia (1993) motivated and defined the third approach, which is our goal in this thesis, by deriving the configuration density of the isotropic Gaussian model, later Díaz-García *et al.* (2003) corrected that result and they also extended it to the case of the central elliptical model, some aspects of those works will be revised here again.

By a revision of the literature around the three above-mentioned methods for statistical shape theory, we find that the least explored approach is the configuration density technique. As we mentioned, there are two main published results concerning that topic: the isotropic Gaussian case proposed by Goodall and Mardia (1993) (and corrected by Díaz-García *et al.* (2003)), and a treatment of the configuration density of Díaz-García *et al.* (2003) for the central elliptical distribution (which resulted invariant under this family).

Moreover, we can identify the following common general facts of the three methods:

1. General size-and-shape, and configuration densities are expanded in terms of zonal polynomials, but at the time of the first appearances of these distributions, no accurate numerical methods for large degrees of the polynomials were given. This forced the use of approximations of the hypergeometric functions to do the inference; but even by using these approaches, the general inference problem (for the QR and SVD approaches) has been studied in the isotropic case and only with $\sigma \to 0$, see Goodall and Mardia (1993). We recall that inference has not been studied in the configuration density method.

2. So the computation of zonal polynomials gives us the key for studying the inference problem. One attempt is due to Gutiérrez *et al.* (2000) which partially solves the numerical computation problem of zonal polynomials, but only recently with the work of Koev and Edelman (2006) efficient computations of the polynomials were given; as a comparison, we read in Koev and Edelman (2006): "We spent (with Gutiérrez *et al.* (2000)) about 8 days to obtain the 627 zonal polynomials degree 20 with a 350 MHz Pentium II processor. In contrast, our Algorithm 4.2 (Koev and Edelman (2006)) takes less than a hundredth of a second to do the same."

   Therefore the use of the algorithms for a numerical computation of zonal polynomials of higher degrees in higher dimensions will let us work on the inference problem.

3. There are two algorithms for hypergeometric functions, based on the

same theoretical source: Koev and Edelman (2006) (numerically) and Dumitriu *et al* (2004) (symbolically). We would like to use directly these routines and avoid the infinite sums of the computed zonal polynomials and the Pochhammer symbols. However, when we contrasted both algorithms for a number of numerical hypergeometric functions we found numerical differences. So before applying these algorithms, we need to check the correct algorithm.

4. Working with a family of distributions (elliptical) will lead to a gamma of unexplored possibilities in the context of shape theory.

5. The preceding works suppose $\sigma^2 \to 0$ (or some complicated constraints for non-small σ^2, see Mardia and Dryden (1989), Dryden and Mardia (1998)) to make inference; it is desirable a theory which does not consider any restriction on the variance.

6. When maximum likelihood estimation is performed and the parameters for shape and variation are estimated simultaneously, classical approaches consider unimodality, see Mardia and Dryden (1989), Dryden and Mardia (1998). We think that the study of the exact distributions via hypergeometric software can consider general (multimodal or unimodal) likelihood.

7. Recall that the object of the Configuration density is to integrate over linear transformations to remove location, scale and uniform shear, and this leads to a confluent hypergeometric form. In this sense we

see a difference with the QR and SVD methods which remove location, rotation and scale, both of them extensively studied in literature.

Now in terms of the configuration density problem we may list the following particular facts:

1. As we said, only the real isotropic Gaussian case has been published. So the whole work of real elliptical configuration densities is waiting to be solved.

2. Even the published works about the isotropic normal configuration density have some details: explicitly, Goodall and Mardia (1993) proposed the density with some errors which were corrected by Díaz-García *et al.* (2003), but the Haar measure employed in the computation of the Jacobian in both papers is not specified (see Muirhead (1982, Section 2.1.4.)), so we need to revise the Gaussian case again.

3. Once the Jacobian is computed we need to study some properties of the function h before the integration over $O(K)$, and $F > 0$. Perhaps new integrals involving zonal polynomials or invariant polynomials of matrix arguments, similar to those presented in Muirhead (1982) and Díaz-García *et al.* (2003), should be solved.

In this thesis we try to supply solutions to the preceding problems, and they are placed and distributed as follows, chapter 1 provides a basic summary on the configuration density problem, the existing results and definitions, and some mathematical elements that we need for finite distribution applications. Then, the main integral which supports all the distributional results and the corresponding corollaries are full derived in chapter 2, and the last section is devoted to the derivation of the configuration density of any elliptical model. Chapter 3 studies the configuration density corresponding to the classical matrix variate elliptical contoured distributions, including, Pearson, Kotz, Bessel, Jensen-Logistic, etc. A summary of chapters 2 and 3 is reported in Caro-Lopera *et al* (2008a). The work ends with chapter 4 which gives a four step procedure for doing inference with the densities here derived and the existing algorithms for computations; then an introduction to the finite configurations is proposed and by using our exact formulae for zonal polynomials, two dimensional applications are studied in Biology (mouse vertebra, gorilla skulls, young craniofacial studies), Medicine (brain MR scans of schizophrenic patients) and image analysis (postcode recognition). A survey of chapter 4 is reported in Caro-Lopera *et al* (2008b).

Finally, the conclusions and subsequent research are written at the end of the thesis.

Chapter 1

Preliminaries

1.1 Shape theory via configuration density

Goodall and Mardia (1993) gives the following motivation for the necessariness of a new shape method. In the rank 1 case, integrating size r of the type ${}_0F_1$ size-and-shape density gives as we expected the type ${}_1F_1$ shape density. This follows from the inductive definition of the hypergeometric functions via the Laplace transform (see Herz (1955)). However, in the matrix setting, the Laplace transform involves integration over the positive definite matrices. Thus, the reflection shape density is not type ${}_1F_1$. Instead, the multivariate approach leads to distributions of type ${}_1F_1$ on equivalence classes of figures modulo affine transformations. The next step would be to type ${}_2F_1$ distributions of canonical correlations, James (1964). See figure 1.1.

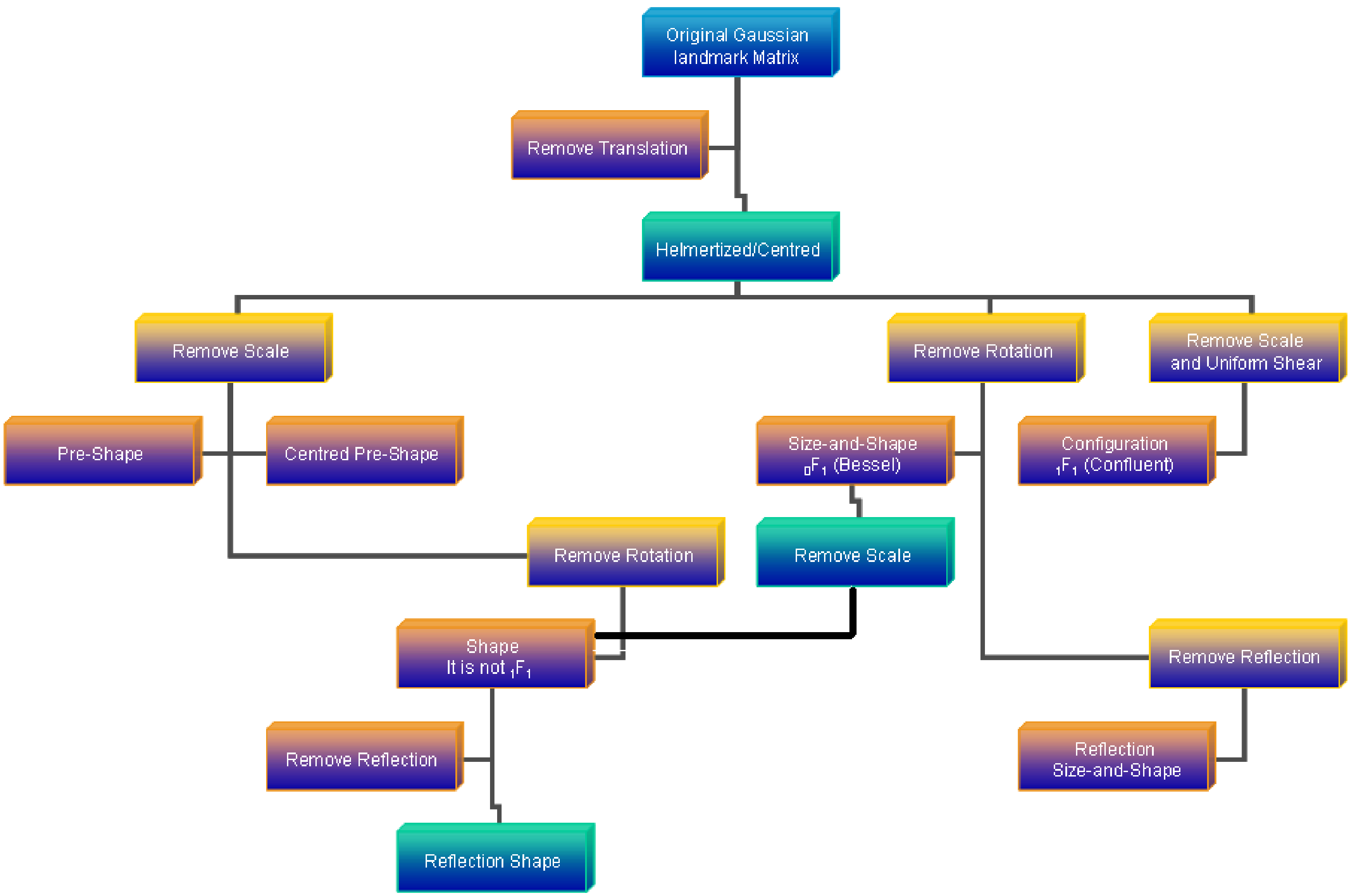

Figure 1.1: Motivation for Configuration Density

Now, an *affine transformation* is specified by a pair $(e : K\times 1, E : K\times K)$ where e is the translation and E is nonsingular. Then we call an equivalence class a *configuration*.

Definition 1.1.1. Two figures $X_1 : N \times K$ and $X_2 : N \times K$ have the same *configuration* if

$$X_2 = X_1 E + 1_N e', \tag{1.1}$$

for some (e, E).

Analogous to the QR decomposition of the centered figure matrix Y, the *configuration coordinates* are constructed in two steps summarized in the expression

$$LX = Y = \begin{pmatrix} I_K \\ V \end{pmatrix} E = UE, \tag{1.2}$$

where $E : K \times K$, $V : q \times K$, $q = N - K - 1$. The matrix $U : N-1 \times K$ contains the configuration coordinates of X, analogous to the size-and-shape coordinates of matrix T in the QR decomposition case. Let $Y_1 : K \times K$ be nonsingular and $Y_2 : q = N - K - 1 \geq 1 \times K$, such that $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$, then $V = Y_2 Y_1^{-1}$ and $E = Y_1$.

As usual, the first common step to the three methods (QR, SVD, configuration) filters the translation from X and it can be achieved by considering contrasts of the data. i.e. pre-multiplying X by a suitable matrix L, for example a Helmert sub-matrix. The Helmert sub-matrix L is the $(N-1)\times N$ Helmert matrix without the first row. The full Helmert matrix R, is a square

$N \times N$ orthogonal matrix with its first row of elements equal to $1/\sqrt{N}$, and the remaining rows are orthogonal to the first row. We drop the first row of R so that the transformed LX does not depend on the original location of the configuration.

Definition 1.1.2. The j-th row of the Helmert sub-matrix L is given by

$$(l_j, \ldots, l_j, -jl_j, 0, \ldots, 0),\ l_j = -[j(j+1)]^{-1/2}, \tag{1.3}$$

and so the j-th row consists of l_j repeated j times, followed by $-jl_j$, and then $N-j-1$ zeros, $j = 1, \ldots, N-1$.

Note that

$$Y = LX \in \mathbb{R}^{K(N-1)} \setminus 0,$$

the origin is removed because coincident landmarks are not allowed. The matrix Y which is refer as the Helmertized landmarks, will be the starting point of any subsequent procedure here; i.e. for distributions and inference, we will always assume that the data is Helmertized.

The centered landmarks are an alternative choice for removing location and are given by

$$X_C = CX,$$

where the centring matrix C is obtained by

$$C = I_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N' = L'L,$$

where $\mathbf{1}_N$ is the $N \times 1$ vector of ones.

We can revert back to the centred landmarks from the Helmertized landmarks by pre-multiplying by L', as

$$L'Y = L'LX = CX.$$

Note that the configuration V is trivial when $N-1 \leq K$, corresponding to 3 or fewer landmarks in $\mathbb{R}^2$, 4 of fewer landmarks in $\mathbb{R}^3$, etc. Without loss of generality, we assume $N-1 > K$, so that $q = N-1-K \geq 1$.

Goodall and Mardia (1993) established the configuration density of the isotropic Gaussian model, however Díaz-García *et al.* (2003) correct the expression as follows

Theorem 1.1.1. *If rank $Y_1 = K$, the configuration density is given by*

$$\frac{2^K \Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{K(N-K-1)/2}\Gamma_K\left(\frac{K}{2}\right)|I+V'V|^{(N-1)/2}} \operatorname{etr}\left[\frac{1}{2\sigma^2}(\mu'U(U'U)^{-1}U'\mu - \mu'\mu)\right]$$
$$ {}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{\mu'U(U'U)^{-1}U'\mu}{2\sigma^2}\right). \tag{1.4}$$

The factor 2^K comes from a non-explicit computation of the jacobian, so we need to revise again that density, and for the correctness of our results, we will derive it by three different methods.

Recall that the expression ${}_1F_1$ in the above density is a polynomial in the latent roots of $\mu'U(U'U)^{-1}U'\mu$.

Finally, Díaz-García *et al.* (2003) extended the above result to the case of the central elliptical model:

Theorem 1.1.2. *Let $Y \sim \mathcal{E}_{N-1\times K}(0, \sigma^2 I_{N-1}, I_K, h)$, then the configuration* *density is invariant under the elliptical family, and is given by*

$$\frac{2^K \Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{K(N-K-1)/2}\Gamma_K\left(\frac{K}{2}\right)}|I + V'V|^{-(N-1)/2}. \tag{1.5}$$

1.2 Zonal polynomials

Now we give a little survey, taken from Muirhead (1982), of a profuse studied theory, the zonal polynomials.

The zonal polynomials of a matrix argument are defined in terms of partitions of positive integers. Let k be a positive integer; a partition κ of k is written as $\kappa = (k_1, k_2, \ldots)$, where $\sum_i k_i = k$, $k_1 \geq k_2 \geq \cdots$, and $k_1, k_2, \ldots$ are not negative integers. We will order the partitions of k lexicographically; i.e, if $\kappa = (k_1, k_2, \ldots)$ and $\lambda = (l_1, l_2, \ldots)$ are two partitions of k we will write $\kappa > \lambda$ if $k_i > l_i$ for the first index for the parts are unequal; some related results on partitions are explored, for example in Caro-Lopera *et al* (2008).

Now, suppose that $\kappa = (k_1, k_2, \ldots)$ and $\lambda = (l_1, l_2, \ldots)$ are two partitions of k (some of the parts may be zero) and let $y_1, \ldots, y_m$ be m variables. If $\kappa > \lambda$ we will say that the monomial $y_1^{k_1} \cdots y_m^{k_m}$ is of higher weight than the monomial $y_1^{l_1} \cdots y_m^{l_m}$.

Then the zonal polynomials can be defined as follows.

Let Y be an $m \times m$ symmetric matrix with latent roots $y_1, \ldots, y_m$ and let $\kappa = (k_1, \ldots, k_m)$ be a partition of k into not more than m parts. The zonal polynomials of Y corresponding to κ, denoted by $C_\kappa(Y)$, is a symmetric, homogeneous polynomial of degree k in the latent roots $y_1, \ldots, y_m$ such that:

- The term of highest weight in $C_\kappa(Y)$ is $y_1^{k_1}\cdots y_m^{k_m}$; that is,

$$C_\kappa(Y) = d_\kappa y_1^{k_1}\cdots y_m^{k_m} + \text{terms of lower weight}, \tag{1.6}$$

where d_κ is a constant.

- $C_\kappa(Y)$ is an eigenfunction of the differential operator Δ_Y given by

$$\Delta_Y = \sum_{i=1}^{m} y_i^2 \frac{\partial^2}{\partial y_i^2} + \sum_{i=1}^{m} \sum_{j=1, j\neq i}^{m} \frac{y_i^2}{y_i - y_j}\frac{\partial}{\partial y_i}. \tag{1.7}$$

- As κ varies over all the partitions of k the zonal polynomials have unit coefficients in the expansions of $(\operatorname{tr} Y)^k$; that is,

$$(\operatorname{tr} Y)^k = (y_1 + \cdots + y_m)^k = \sum_\kappa C_\kappa(Y). \tag{1.8}$$

It can be proved that the zonal polynomial $C_\kappa(Y)$ satisfies the following partial differential equation

$$\Delta_Y C_\kappa(Y) = [k(m-1) + \sum_{i=1}^{m} k_i(k_i - i)] C_\kappa(Y). \tag{1.9}$$

The zonal polynomials are expressed in terms of a number of basis, one of them are the monomial symmetric functions. If $k_1, \ldots, k_m$, the monomial symmetric function of $y_1, \ldots, y_m$ corresponding to κ is defined as

$$M_\kappa(Y) = \sum \cdots \sum y_{i_1}^{k_1} y_{i_2}^{k_2} \cdots y_{i_p}^{k_p}, \tag{1.10}$$

where p is the number of nonzero parts in the partition κ and the summation is over the distinct permutations $(i_1, \ldots, i_p)$ of p different integers from the integers $1, \ldots, m$. Hence

$$M_\kappa(Y) = y_1^{k_1}\cdots y_m^{k_m} + \text{symmetric terms}. \tag{1.11}$$

The differential equation for $C_\kappa(Y)$ gives rise to a recurrence relation between the coefficients of the monomial symmetric function in C_κ; once the coefficient of the term of highest weight is given, the other coefficients are uniquely determined by the recurrence relation. So the zonal polynomials can be expressed in terms of the monomial symmetric functions as

$$C_\kappa(Y) = \sum_{\lambda \leq \kappa} c_{\kappa,\lambda} M_\lambda(Y), \tag{1.12}$$

and this leads to the following recurrence relation for the coefficients,

$$c_{\kappa,\lambda} = \sum_{\lambda < \mu \leq \kappa} \frac{[(l_i + t) - (l_j - t)]}{\rho_\kappa - \rho_\lambda} c_{\kappa,\mu}, \tag{1.13}$$

where $\lambda = (l_1, \ldots, l_m)$ and $\mu = (l_1, \ldots, l_i + t, \ldots, l_j - t, \ldots, l_m)$ for $t = 1, \ldots, l_j$ such that, when the parts of the partition μ are arrange in descending order then $\lambda < \mu \leq \kappa$. This algorithm is the base for the fast routines by Koev and Edelman (2006).

However, no general formula for zonal polynomials is known; only the above partial differential equation has been solved for $m = 2$, (James (1968)), which is a particular case of a more general result established for Jack polynomials by Caro-Lopera *et al* (2007). This formulae are useful for computing finite configuration densities for two dimensional applications, as we shall see in the last chapter of this thesis. Formulae for third degree (valid for three dimensional applications) are only known in a recurrence way, see for example James (1964). However, the algorithms by Koev and Edelman (2006) can be applied for any infinite configuration density of any dimension under certain truncation assumptions.

The above results hold for zonal polynomials of positive definite matrix argument, however they can be studied in the semidefinite positive case, see Díaz-García *and* Caro-Lopera (2006).

Now we give some summary of Jack polynomials and exact two dimensional formulae, which particularized to zonal polynomials, will be needed in finite configuration applications.

Let us characterize the Jack symmetric function $J_{\kappa}^{(\alpha)}(y_1, \ldots, y_m)$ of parameter α, see Sawyer (1997). As we said before, a decreasing sequence of nonnegative integers $\kappa = (k_1, k_2, \ldots)$ with only finitely many nonzero terms is said to be a partition of $k = \sum k_i$. Let κ and $\lambda = (l_1, l_2, \ldots)$ be two partitions of k. We write $\lambda \leq \kappa$ if $\sum_{i=1}^{t} l_i \leq \sum_{i=1}^{t} k_i$ for each t. The conjugate of κ is $\kappa' = (k_1', k_2', \ldots)$ where $k_i' = \text{card}\{j : k_j \geq i\}$. The length of κ is $l(k) = \max\{i : k_i \neq 0\} = k_1'$. If $l(\kappa) \leq m$, one often writes $\kappa = (k_1, k_2, \ldots, k_m)$. The partition $(1, \ldots, 1)$ of length m will be denoted by 1_m.
And recall that the monomial symmetric function $M_{\kappa}(\cdot)$ indexed by a partition κ can be regarded as a function of an arbitrary number of variables such that all but a finite number are equal to 0: if $y_i = 0$ for $i > m \geq l(\kappa)$ then $M_{\kappa}(y_1, \ldots, y_m) = \sum y_1^{\sigma_1} \cdots y_m^{\sigma_m}$, where the sum is over all distinct permutations $\{\sigma_1, \ldots, \sigma_m\}$ of $\{k_1, \ldots, k_m\}$, and if $l(\kappa) > m$ then $M_{\kappa}(y_1, \ldots, y_m) = 0$. A symmetric function f is a linear combination of monomial symmetric functions. If f is a symmetric function then $f(y_1, \ldots, y_m, 0) = f(y_1, \ldots, y_m)$. For each $m \geq 1$, $f(y_1, \ldots, y_m)$ is a symmetric polynomial in m variables.

Thus the Jack symmetric function $J_{\kappa}^{(\alpha)}(y_1, \ldots, y_m)$ with a parameter α, satisfy the following conditions:

$$J_\kappa^{(\alpha)}(y_1,\ldots,y_m) = \sum_{\lambda\le\kappa} j_{\kappa,\lambda} M_\lambda(y_1,\ldots,y_m), \quad (1.14)$$

$$J_\kappa^{(\alpha)}(1,\ldots,1) = \alpha^k \prod_{i=1}^{m} \left(\frac{m-i+1}{\alpha}\right)_{k_i}, \quad (1.15)$$

$$\sum_{i=1}^{m} y_i^2 \frac{\partial^2 J_\kappa^{(\alpha)}(y_1,\ldots,y_m)}{\partial y_i^2} + \frac{2}{\alpha}\sum_{i=1}^{m} y_i^2 \sum_{j\ne i} \frac{1}{y_i-y_j}\frac{\partial J_\kappa^{(\alpha)}(y_1,\ldots,y_m)}{\partial y_i} = \sum_{i=1}^{m} k_i(k_i-1+\frac{2}{\alpha}(m-i))J_\kappa^{(\alpha)}(y_1,\ldots,y_m). \quad (1.16)$$

Here the constants $j_{\kappa,\lambda}$ do not dependent on $y_i's$ but on κ and λ, and $(a)_n = \prod_{i=1}^n (a+i-1)$. Note that if $m < l(\kappa)$ then $J_\kappa^{(\alpha)}(y_1,\ldots,y_m) = 0$. The conditions include the case $\alpha = 0$ and then $J_\kappa^{(0)}(y_1,\ldots,y_m) = e_{\kappa'}\prod_{i=1}^m (m-i+1)^{k_i}$, where $e_\kappa(y_1,\ldots,y_m) = \prod_{i=1}^{l(\kappa)} e_{k_i}(y_1,\ldots,y_m)$ are the elementary symmetric functions indexed by partitions κ, if $m \ge l(\kappa)$ then $e_r(y_1,\ldots,y_m) = \sum_{i_1<i_2<\cdots<i_r} y_{i_1}\cdots y_{i_r}$, and if $m < l(\kappa)$ then $e_r(y_1,\ldots,y_m) = 0$, see Sawyer (1997).

Now, from Koev and Edelman (2006), the Jack functions $J_\kappa^{(\alpha)}(Y) = J_\kappa^{(\alpha)}(y_1,\ldots,y_m)$, with $y_1,\ldots,y_m$ being the eigenvalues of the matrix Y, can be normalized in such a way that

$$\sum_\kappa C_\kappa^\alpha(Y) = (\mathrm{tr}(Y))^k,$$

where $C_\kappa^\alpha(Y)$ denotes the Jack polynomials. They are related to the Jack functions by

$$C_\kappa^\alpha(Y) = \frac{\alpha^k k!}{j_\kappa} J_\kappa^\alpha(Y), \quad (1.17)$$

where

$$j_\kappa = \prod_{(i,j)\in\kappa} h_*^\kappa(i,j) h_\kappa^*(i,j),$$

and $h_*^\kappa(i,j) = k_j - i + \alpha(k_i - j + 1)$ and $h_\kappa^*(i,j) = k_j - i + 1 + \alpha(k_i - j)$ are the upper and lower hook lengths at $(i,j) \in \kappa$, respectively.

Then by applying (1.17), we can write (1.16) as

$$\sum_1^m y_i^2 \frac{\partial^2 C_\kappa^{(\alpha)}(Y)}{\partial y_i^2} + \frac{2}{\alpha} \sum_{i=1}^m y_i^2 \sum_{j\neq i} \frac{1}{y_i - y_j} \frac{\partial C_\kappa^{(\alpha)}(Y)}{\partial y_i} =$$

$$\sum_{i=1}^m k_i(k_i - 1 + \frac{2}{\alpha}(m-i)) C_\kappa^{(\alpha)}(Y). \quad (1.18)$$

Now, when $m = 2$ in (1.18), Caro-Lopera *et al* (2007) found the following formulae for Jack Polynomials of the Second Order

$$\frac{C_{(k_1,k_2)}^{(\alpha)}(Y)}{C_{(k_1,k_2)}^{(\alpha)}(I_2)} = (y_1 y_2)^{(k_1+k_2)/2} A_1 F\left(-\frac{\rho}{2}, \frac{\rho}{2} + \frac{1}{\alpha}; \frac{1}{2}; \frac{(y_1+y_2)^2}{4y_1y_2}\right)$$

$$+ \frac{(y_1y_2)^{(k_1+k_2-1)/2}}{2(y_1+y_2)^{-1}} A_2 F\left(\frac{1}{\alpha} + \frac{1+\rho}{2}, \frac{1}{2} - \frac{\rho}{2}; \frac{3}{2}; \frac{(y_1+y_2)^2}{4y_1y_2}\right), \quad (1.19)$$

with ρ being either even or odd. For distinguishing the case under consideration, odd or even, we will use the upper indices o or e with A_1 and A_2. Then the corresponding solutions are the following

Even case. If $\rho = k_1 - k_2 = 2n$, $n = 0, 1, 2, \ldots$ then

$$A_1^e = \frac{(-1)^n \prod_{i=0}^{n-1}(1+2i)}{\prod_{i=0}^{n-1}\left(1 + 2\left(\frac{1}{\alpha} + i\right)\right)} \quad \text{and} \quad A_2^e = 0.$$

Odd case. If $\rho = k_1 - k_2 = 2n+1$, $n = 0, 1, 2, \ldots$ then

$$A_1^o = 0 \quad \text{and} \quad A_2^o = (2n+1)A_1^e.$$

Three particular cases are of interest in the literature: the quaternionic case ($\alpha = 1/2$), the complex zonal polynomials ($\alpha = 1$) and the real zonal polynomials ($\alpha = 2$), these results are summarized in the following table:

α	ρ	a	b	c	A_1	A_2
$\frac{1}{2}$	even	$-n$	$n+2$	$\frac{1}{2}$	$\frac{(-1)^n 3}{(2n+1)(2n+3)}$	0
	odd	$n+3$	$-n$	$\frac{3}{2}$	0	$\frac{(-1)^n 3}{(2n+3)}$
1	even	$-n$	$n+1$	$\frac{1}{2}$	$\frac{(-1)^n}{(2n+1)}$	0
	odd	$n+2$	$-n$	$\frac{3}{2}$	0	$(-1)^n$
2	even	$-n$	$n+1/2$	$\frac{1}{2}$	$\frac{(-1)^n (2n)!}{2^{2n}(n!)^2}$	0
	odd	$n+3/2$	$-n$	$\frac{3}{2}$	0	$\frac{(-1)^n (2n+1)!}{2^{2n}(n!)^2}$

The above formula for the real zonal polynomials corresponds to that derived by James (1968) and for the complex zonal polynomials, obtained by Caro-Lopera *et al* (2006). This is the first appearance of an exact formulae for quaternionic polynomials.

We end this section with a concept (see Muirhead (1982)) which will be useful in the Gaussian configuration density context.

The hypergeometric functions of matrix arguments are given by

$$ {}_pF_q(a_1,\ldots,a_p;b_1,\ldots,b_q;X)=\sum_{k=0}^{\infty}\frac{1}{k!}\sum_{\kappa}\frac{(a_1)_\kappa\cdots(a_p)_\kappa}{(a_1)_\kappa\cdots(a_q)_\kappa}C_\kappa(X), \tag{1.20}$$

where $\sum_\kappa$ denotes summation over all partitions $\kappa(k_1,\ldots,k_m)$, $k_1\geq\ldots k_m\geq 0$ of k, and

$$(a)_\kappa=\prod_{i=1}^{m}\left(a-\frac{1}{2}(i-1)\right)_{k_i}, \tag{1.21}$$

where

$$(a)_k=a(a+1)\cdots(a+k-1),\ (a)_0=1. \tag{1.22}$$

Here X, the argument of the function, is a complex symmetric $m\times m$ matrix and the parameters a_i,b_j are arbitrary complex. No denominator parameter b_j is allowed to be zero or an integer of half-integer $\leq \frac{1}{2}(m-1)$ (otherwise some of the denominators in the series will vanish). If any numerator parameter a_i is negative integer, say, $a_1=-n$, then the function is a polynomial of degree mn, because $k\geq mn+1$, $(a_1)_\kappa=(-n)_\kappa=0$. The series converges for all X if $p\leq q$, it converges for $\|X\|<1$ if $p=q+1$, $\|X\|$ denotes the maximum of the absolute values of the latent roots of X, and unless it terminates, it diverges for $X\neq 0$ if $p>q+1$. Finally, when $m=1$ the series reduces to the classical hypergeometric function.

Some important cases are

$$ {}_0F_0(X) = \mathrm{etr}(X). \tag{1.23} $$

$$ {}_1F_0(a; X) = |I_m - X|^{-a}, \quad \|X\| < 1. \tag{1.24} $$

If X is an $m \times n$ real matrix with $m \leq n$ and $H = [H_1 : H_2] \in O(n)$, where H_1 is $n \times m$ then

$$ {}_0F_1(\frac{1}{2}n; \frac{1}{4}XX') = \int_{O(n)} \mathrm{etr}(XH_1)(dH), \tag{1.25} $$

where (dH) denotes de normalized invariant measure on $O(n)$.

$$ \begin{aligned} {}_1F_1(a; c; X) &= \frac{\Gamma_m(c)}{\Gamma(a)\Gamma(c-a)} \int_{0<Y<I_m} \mathrm{etr}(XY)|Y|^{a-(m+1)/2} \\ &\quad |I - Y|^{c-a-(m+1)/2}(dY), \end{aligned} \tag{1.26} $$

valid for all symmetric X, and $\mathrm{Re}(a), \mathrm{Re}(c), \mathrm{Re}(c-a) > \frac{1}{2}(m-1)$.

$$ \begin{aligned} {}_2F_1(a, b; c; X) &= \frac{\Gamma_m(c)}{\Gamma(a)\Gamma(c-a)} \int_{0<Y<I_m} |I - XY|^{-b}|Y|^{a-(m+1)/2} \\ &\quad |I - Y|^{c-a-(m+1)/2}(dY), \end{aligned} \tag{1.27} $$

valid for $\mathrm{Re}(X) < I$, and $\mathrm{Re}(a), \mathrm{Re}(c-a) > \frac{1}{2}(m-1)$.

And Kummer and Euler relations, respectively (Herz (1955))

$$ {}_1F_1(a; c; X) = \mathrm{etr}(X)\, {}_1F_1(c-a; c; -X). \tag{1.28} $$

$$ \begin{aligned} {}_2F_1(a, b; c; X) &= |I - X|^{-b}\, {}_2F_1(c-a, b; c; -X(I-X)^{-1}) \\ &= |I - X|^{c-a-b}\, {}_2F_1(c-a, c-b; c; X). \end{aligned} \tag{1.29} $$

1.3 Elliptically contoured distributions

The elliptically contoured distribution of a random matrix has been studied by various authors, including Fang and Zhang (1990) and Gupta and Varga (1993).

Here we just give their definition. We say that $X : N \times K$ has a matrix variate elliptically contoured distribution if its density respect to the Lebesgue measure is given by:

$$f_X(X) = \frac{1}{|\Sigma|^{K/2}|\Theta|^{N/2}} h(\operatorname{tr}((X-\mu)'\Sigma^{-1}(X-\mu)\Theta^{-1})), \tag{1.30}$$

where $\mu : N \times K$, $\Sigma : N \times N$, $\Theta : K \times K$, Σ positive definite ($\Sigma > 0$), $\Theta > 0$. Such a distribution is denoted by $X \sim \mathcal{E}_{N\times K}(\mu, \Sigma, \Theta, h)$. We refer in this work to (1.30) which we always assume that it exists.

The classical families of elliptically contoured distributions are given next.

1.3.1 Matrix variate Kotz type distribution

The $p \times n$ random matrix X is said to have a matrix variate symmetric Kotz type distribution with parameters $T, R, s \in \Re$, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $R > 0$, $s > 0$ $2T + np > 2$, $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{sR^{\frac{2T+np-2}{2s}}\Gamma\left(\frac{np}{2}\right)}{\pi^{np/2}\Gamma\left(\frac{2T+np-2}{2s}\right)|\Sigma|^{n/2}|\Phi|^{p/2}}\left[\operatorname{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right]^{T-1}$$
$$\exp\left\{-R\operatorname{tr}\left[(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right]^{s}\right\}.$$

When $T = s = 1$, and $R = 1/2$ we get the probability density function of the absolutely continuous matrix variate normal distribution.

1.3.2 Matrix variate normal distribution

The $p \times n$ random matrix X is said to have a matrix variate normal distribution with parameters, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}|\Phi|^{p/2}}\operatorname{etr}\left[-\frac{1}{2}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right].$$

1.3.3 Matrix variate Pearson type VII distribution

A $p \times n$ random matrix X is said to have a matrix variate symmetric Pearson type VII distribution with parameters $s, R \in \mathbb{R}$, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $R > 0$, $s > np/2$, $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{\Gamma(s)}{(\pi R)^{np/2}\Gamma\left(s - \frac{np}{2}\right)|\Sigma|^{n/2}|\Phi|^{p/2}}\left(1 + \frac{\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}}{R}\right)^{-s}.$$

When $s = (np + R)/2$, X is said to have a matrix variate t-distribution with R degrees of freedom. And in this case, if $R = 1$, then X is said to have a matrix variate Cauchy distribution.

1.3.4 Matrix variate Bessel distribution

Another interesting elliptical distribution is the so called Bessel distribution, explicitly, the $p \times n$ random matrix is said to have a matrix variate symmetric Bessel distribution with parameters $q, r \in \mathbb{R}$, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $r > 0$, $q > -\frac{np}{2}$, $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{[\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}]^{\frac{q}{2}} K_q\left(\frac{[\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}]^{\frac{1}{2}}}{r}\right)}{2^{q+np-1}\pi^{\frac{np}{2}} r^{np+q}\Gamma\left(q + \frac{np}{2}\right)|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}},$$

where $K_q(z)$ is the modified Bessel function of the third kind; that is

$$K_q(z) = \frac{\pi}{2}\frac{I_{-q}(z) - I_q(z)}{\sin(q\pi)}, \quad |\arg(z)| < \pi, \quad q \text{ is integer},$$

and

$$I_q(z) = \sum_{k=0}^{\infty}\frac{1}{k!\Gamma(k+q+1)}\left(\frac{z}{2}\right)^{q+2k}, \; |z| < \infty, |\arg(z)| < \pi.$$

If $q = 0$ and $r = \frac{\sigma}{\sqrt{2}}$, $\sigma > 0$, this distribution is known as the matrix variate Laplace distribution.

1.3.5 Matrix variate Jensen-Logistic distribution

This density was proposed by Jensen (1985) and studied by Fang *et al* (1990), Gupta and Varga (1993) among many others.

Note that this distribution is not a generalization of the classical univariate logistic, thus we will refer it in the thesis as the matrix variate *Jensen-Logistic* distribution.

In this case we say that the $p \times n$ random matrix X is has a matrix variate symmetric Jensen-logistic distribution with parameters $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{c \operatorname{etr} -(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}}{|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}\left(1+\operatorname{etr} -(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right)^2},$$

where

$$c = \frac{\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)}\int_0^\infty z^{\frac{np}{2}-1}\frac{\mathrm{e}^{-z}}{(1+\mathrm{e}^{-z})^2}dz,$$

see Jensen (1985), Fang *et al* (1990) and Gupta and Varga (1993).

Chapter 2

Configuration Density

In this chapter we provide the technical tools for finding the configuration density under an elliptical model. With the main integral, a sort of known results of the matrix multivariate analysis literature, are straightforward derived without the classical multiple technique, instead of that, they are obtained by computing a single integral; then some errors in literature are detected and corrected. Once the main integral is established, the non-isotropic noncentral elliptical configuration density is derived, and the isotropic central and the isotropic gaussian cases (the only published works in this area) are corrected and presented as corollaries.

2.1 The main integral

Matrix generalization of elliptically contoured distributions, and the respective generalization to the classical multivariate matrix variate theory based on normality, has propitiated a number of results which cover a gamma of new distributions. However, integration over positive definite symmetric spaces involving zonal polynomials, remains problematic even in the normal multivariate case, widely studied by Muirhead (1982), for example. So the elliptical generalization keeps these difficulties.

Noncentral elliptical multivariate distributions involve a number of general integrals to study, it depends on the transformations under consideration, but all of them are founded in an important fact, the elliptically contoured distribution are characterized by a symmetric function, say, $h(U)$, i.e, $h(AB) = h(BA)$, for any squared matrix A and B. The simplest function we can consider is the trace of a positive definite matrix and then the zonal polynomials arise naturally, then euclidian and affine transformations of the random matrix will precise of integration over the orthogonal group and the positive definite space, etc. In the case of positive definite matrices, we find in Constantine (1963) the source for all the posterior works, in fact it inspired the following general result for elliptical integration (see Xu and Fang (1989):

Lemma 2.1.1.

$$\begin{aligned}\int_{X>0} \quad & h(XZ)|X|^{a-(m+1)/2}C_{\kappa}(XY)(dX) \\ &= \frac{|Z|^{-a}C_{\kappa}(YZ^{-1})}{C_{\kappa}(I_m)}\int_{X>0} h(X)|X|^{a-(m+1)/2}C_{\kappa}(X)(dX), \qquad (2.1)\end{aligned}$$

where Y is a symmetric $m \times m$ matrix,Z is a complex symmetric $m \times m$ matrix, $Re(Z) > 0$ and $C_{\kappa}(X)$ is the zonal polynomial of X, see Muirhead (1982).

As a convention, in this work we always assume that the integrals we meet with exist.

Only a few number of functions $h(\cdot)$ have been studied, they constituted the works of Herz, James, Constantine and Khatri, among many others, and an excellent survey of this works and their statistical applications are given by Muirhead, for example. We can check easily that all the publish works for particular $h(\cdot)$'s depend on the following result due to Constantine (1963) (motivated by Littlewood (1950)):

Lemma 2.1.2. *If* $Y = \mathrm{diag}(y_1, \ldots, y_m)$ *and* $X = (x_{ij})$ *is an* $m \times m$ *positive definite matrix then*

$$\begin{aligned} C_\kappa(XY) &= d_\kappa y_1^{k_1} \cdots y_m^{k_m} |X_1|^{k_1-k_2} |X_2|^{k_2-k_3} \cdots |X|^{k_m} \\ &\quad + \text{terms of lower weight in the } y\text{'s}, \end{aligned} \tag{2.2}$$

where $X_p = (x_{ij})$, $i, j = 1, \ldots p$, $\kappa = (k_1, \ldots, k_m)$ *and* d_κ *is the coefficient of the term of highest weight in* $C_\kappa(\cdot)$, *see Muirhead (1982), p.228.*

If Y is replaced by Y^{-1}, the above result turns

Corollary 2.1.1. *If* $Y = \mathrm{diag}(y_1, \ldots, y_m)$ *and* $X = (x_{ij})$ *is an* $m \times m$ *positive definite matrix then*

$$\begin{aligned} C_\kappa(XY^{-1}) &= d_\kappa y_1^{k_1} \cdots y_m^{k_m} |X_1|^{-(k_1-k_2)} |X_2|^{-(k_2-k_3)} \cdots |X|^{-k_m} \\ &\quad + \text{terms of lower weight in the } y\text{'s}, \end{aligned} \tag{2.3}$$

where $X_p = (x_{ij})$, $i, j = 1, \ldots p$, $\kappa = (k_1, \ldots, k_m)$ *and* d_κ *is the coefficient of the term of highest weight in* $C_\kappa(\cdot)$, *see Muirhead (1982), p.256.*

The main obstacle for getting new integrals (and then new statistical applications) comes from the multiple integration on $X > 0$ given in lemma 2.1.1; we will see that the trace version of the multivariate integral can be reduced to the computation of a single integral, if it exists. The following main integral can be seen as a combination of two separated results by Xu and Fang (1989) and

Teng *et al* (1989). However we prove it, with a didactical purpose, by using the same standard procedure of Constantine (1963) and the multiple integral simplification given by, for example, Caro *and* Nagar (2006) or Fang and Zhang (1990).

Theorem 2.1.1. *Let Z be a complex symmetric $m \times m$ matrix with $Re(Z) > 0$ and let Y be a symmetric $m \times m$ matrix. Then*

$$\int_{X>0} h(\operatorname{tr} XZ)|X|^{a-(m+1)/2} C_\kappa(XY)(dX) = \frac{|Z|^{-a}(a)_\kappa \Gamma_m(a) C_\kappa(YZ^{-1})}{\Gamma(ma+k)} S, \tag{2.4}$$

where

$$S = \int_0^\infty h(w) w^{ma+k-1} dw < \infty, \tag{2.5}$$

$$(a)_\kappa = \prod_{i=1}^m \left(a - \frac{1}{2}(i-1)\right)_{k_i}, \ \kappa = (k_1, \ldots, k_m),\ k_1 \geq \cdots \geq k_m > 0, \sum_{i=1}^m k_i = k,$$
$$(a)_k = a(a+1)\cdots(a+k-1),\ (a)_0 = 1,$$

and

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma\left[a - \frac{1}{2}(i-1)\right], Re(a) > \frac{1}{2}(m-1).$$

Proof. The first part related to the normalization procedure comes from lemma 2.1.1 by replacing $h(\cdot)$ with $h(\operatorname{tr}\cdot)$, however for a didactic proposal we give the main details.

Denote the left hand side of (2.4) by $g(Y, Z)$, then

$$g(Y, I_m) = \int_{X>0} h(\operatorname{tr} X)|X|^{a-(m+1)/2} C_\kappa(XY)(dX). \tag{2.6}$$

For any $H \in O(m)$ we have:

$$g(HYH', I_m) = \int_{X>0} h(\operatorname{tr} X)|X|^{a-(m+1)/2} C_\kappa(XHYH')(dX). \tag{2.7}$$

Let $U = H'XH$, which implies $(dU) = (dX)$, so

$$\begin{aligned} g(HYH', I_m) &= \int_{U>0} h^s(\mathrm{tr}(-U))|U|^{a-(m+1)/2} C_\kappa(UY)(dU) \\ &= g(Y, I_m). \end{aligned} \tag{2.8}$$

Thus $g(Y, I_m)$ is a symmetric function of Y. Because of (2.7) and (2.8) we get on integrating with respect to the normalized invariant measure (dH) on $O(K)$ that:

$$\begin{aligned} g(Y, I_m) &= g(Y, I_m)1 \\ &= g(Y, I_m)\int_{O(K)} (dH) \\ &= \int_{O(K)} g(Y, I_m)(dH) \\ &= \int_{O(K)} g(HYH', I_m)(dH) \\ &= \int_{O(K)}\int_{X>0} h(\mathrm{tr}\, X)|X|^{a-(m+1)/2} C_\kappa(XHYH')(dH)(dX) \\ &= \int_{X>0} h(\mathrm{tr}\, X)|X|^{a-(m+1)/2}\int_{O(K)} C_\kappa(XHYH')(dH)(dX) \\ &= \int_{X>0} h(\mathrm{tr}\, X)|X|^{a-(m+1)/2}\frac{C_\kappa(X)C_\kappa(Y)}{C_\kappa(I_m)}(dX), \\ &= \frac{C_\kappa(Y)}{C_\kappa(I_m)}\int_{X>0} h(\mathrm{tr}\, X)|X|^{a-(m+1)/2} C_\kappa(X)(dX) \\ &= \frac{g(I_m, I_m)}{C_\kappa(I_m)} C_\kappa(Y), \end{aligned} \tag{2.9}$$

where James (1964) was used for the integration over $O(m)$; which ratifies (2.1) for $Z = I_m$ and this particular h, (up here the mentioned replicated procedure of Xu and Fang (1989) which only normalized the original integral).

Now for reducing the multiple integral in (2.9) consider the following facts.

By definition of zonal polynomials

$$\begin{aligned} g(Y, I_m) &= \frac{g(I_m, I_m)}{C_\kappa(I_m)} C_\kappa(Y) \\ &= \frac{g(I_m, I_m)}{C_\kappa(I_m)} d_\kappa y_1^{k_1} \cdots y_m^{k_m} + \text{terms of lower weight in the } y\text{'s}, \end{aligned} \tag{2.10}$$

and by lemma 2.1.2

$$\begin{aligned} & g(Y, I_m) = \int_{X>0} h(\operatorname{tr} X)|X|^{a-(m+1)/2} C_\kappa(XY)(dX) \\ & = d_\kappa y_1^{k_1} \cdots y_m^{k_m} \int_{X>0} h(\operatorname{tr} X)|X|^{a-(m+1)/2} |X_1|^{k_1-k_2} |X_2|^{k_2-k_3} \cdots |X|^{k_m}(dX) \\ & + \text{terms of lower weight in the } y\text{'s}. \end{aligned} \tag{2.11}$$

To evaluate the last integral, put $X = T'T$, where T is upper triangular with positive diagonal elements. Then

$$\operatorname{tr} X = \sum_{i \le j}^{m} t_{ij}^2,\ |X_1| = t_{11}^2,\ |X_2| = t_{11}^2 t_{22}^2, \ldots, |X| = \prod_{i=1}^{m} t_{ii}^2,$$

with jacobian

$$(dX) = 2^m \prod_{i=1}^{m} t_{ii}^{m+1-i} (dT),$$

see Muirhead (1982), p.60.

By using Caro *and* Nagar (2006) or Fang and Zhang (1990), we have

$$\begin{aligned}
g(Y, I_m) &= d_\kappa y_1^{k_1} \cdots y_m^{k_m} \int \cdots \int_{t_{ij}} h\left(\sum_{i\le j}^{m} t_{ij}^2\right) \prod_{i=1}^{m} t_{ii}^{2a+2k_i-i} 2^m (dT) \\
&\quad + \text{terms of lower weight in the } y\text{'s} \\
&= d_\kappa y_1^{k_1} \cdots y_m^{k_m} \frac{\prod_{i=1}^m \Gamma\left(a+k_i-\frac{1}{2}(i-1)\right) \prod_{i<j}^m \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\sum_{i=1}^m \left(a+k_i-\frac{1}{2}(i-1)\right)+\frac{1}{4}m(m-1)\right)} \\
&\quad \int_0^\infty h(w) w^{\sum_{i=1}^m \left(a+k_i-\frac{1}{2}(i-1)\right)+\frac{1}{4}m(m-1)-1} dw \\
&\quad + \text{terms of lower weight in the } y\text{'s} \\
&= d_\kappa y_1^{k_1} \cdots y_m^{k_m} \frac{\prod_{i=1}^m \Gamma\left(a+k_i-\frac{1}{2}(i-1)\right) \pi^{\frac{1}{4}m(m-1)}}{\Gamma(ma+k)} \\
&\quad \int_0^\infty h(w) w^{ma+k-1} dw \\
&\quad + \text{terms of lower weight in the } y\text{'s} \\
&= d_\kappa y_1^{k_1} \cdots y_m^{k_m} \frac{(a)_\kappa \Gamma_m(a)}{\Gamma(ma+k)} S \\
&\quad + \text{terms of lower weight in the } y\text{'s}, \qquad (2.12)
\end{aligned}$$

where the last line follows from Muirhead (1982), p. 248 and we denote

$$S = \int_0^\infty h(w) w^{ma+k-1} dw < \infty.$$

Equating coefficients of $y_1^{k_1} \cdots y_m^{k_m}$ in (2.12) and (2.10) we have that

$$\frac{g(I_m, I_m)}{C_\kappa(I_m)} = \frac{(a)_\kappa \Gamma_m(a)}{\Gamma(ma+k)} S.$$

And using this in (2.9), we obtain

$$\int_{X>0} h(\operatorname{tr} X) |X|^{a-(m+1)/2} C_\kappa(XY)(dX) = \frac{(a)_\kappa \Gamma_m(a)}{\Gamma(ma+k)} C_\kappa(Y) S, \qquad (2.13)$$

which establishes (2.4) for $Z = I_m$; this result was derived by Teng *et al* (1989).

Now consider the integral (2.4) when $Z > 0$ is real and put $V = Z^{1/2} X Z^{1/2}$, so that $(dV) = |Z|^{(m+1)/2}(dX)$, then (2.4) becomes

$$\begin{aligned} g(Y,Z) &= |Z|^{-a}\int_{V>0} h(\operatorname{tr}V)|V|^{a-(m+1)/2}C_\kappa(VZ^{-1/2}YZ^{-1/2})(dV) \\ &= \frac{|Z|^{-a}(a)_\kappa\Gamma_m(a)}{\Gamma(ma+k)}C_\kappa(YZ^{-1})S, \end{aligned}$$

where the last line follows from (2.13). Thus the theorem is true for real $Z > 0$ and it follows for complex Z with $Re(Z) > 0$ by analytic continuation. ■

When the zonal polynomial in (2.4) is evaluated in the latent roots of $X^{-1}Y$ the respective result was obtained by Runze (1997), in fact, that work proposed (without proof) the expression for the expected values of an invariant polynomial under a distribution type of (2.4) and then it was applied in the expectation of zonal polynomials respect to a Pearson VII type distribution; even more, the last particular expectation problem was full studied by Xu and Fang (1989). But, as we shall see in subsection 2.1.1, the first result has an error (perhaps a typographical) and the second and third results are incorrect. In any case, those expressions can be stated easily as corollaries of theorem 2.1.1.

Next we generalize a result of Khatri (1966) (lemma 7, eq. (19)), which was proved by expanding a Laplace transform; here the proof is straightforward.

Corollary 2.1.2.

$$\begin{aligned} &\int_{X>0}\operatorname{etr}(-XZ)(\operatorname{tr}(XZ))^j|X|^{a-(m+1)/2}C_\kappa(XY)(dX) \\ &\qquad = \frac{\Gamma_m(a)(a)_\kappa\Gamma(ma+j+k)}{\Gamma(ma+k)}|Z|^{-a}C_\kappa(YZ^{-1}), \qquad (2.14) \end{aligned}$$

where Y is a symmetric $m \times m$ matrix and Z is a complex symmetric $m \times m$ matrix with $Re(Z) > 0$.

Proof. Take $h(y) = e^{-y}y^j$ in (2.5), then

$S = \int_0^\infty h(y) y^{ma+k-1} dy = \Gamma(ma + j + k)$ and (2.14) follows. When $Z = I_m$ we have lemma 7, eq. (19) of Khatri (1966). ■

The result of Constantine (1963) (eq.(1)), widely used by Muirhead (1982), is obtained in the same way

Corollary 2.1.3.

$$\int_{X>0} \operatorname{etr}(-XZ)|X|^{a-(m+1)/2} C_\kappa(XY)(dX)$$
$$= (a)_\kappa \Gamma_m(a)|Z|^{-a} C_\kappa(YZ^{-1}), \tag{2.15}$$

where Y is a symmetric $m \times m$ matrix and Z is a complex symmetric $m \times m$ matrix with $Re(Z) > 0$.

Proof. Or just replace $j = 0$ in the preceding corollary. ■

We already have the tools for finding the configuration densities of any elliptical model, but first we need to revise some published results for a particular model.

2.1.1 On some Pearson VII type published results

The Pearson VII type version of (2.13) and some applications were full detailed by Xu and Fang (1989) in example 3.2, p.474-476 by the classical multivariate procedure in order to study the generalized Wishart matrix produced by a Pearson VII type distribution; however the same integral give us the Pearson VII type configuration density, but after comparisons with those works we find important differences, then we must to revise them again and for the correctness of our results we write this subsection and derive again the results by different ways.

It is straightforward that

Corollary 2.1.4.

$$\int_{W>0}\left(1+m^{-1}\operatorname{tr}W\right)^{-(m+np)/2}|W|^{\frac{n}{2}-\frac{(p+1)}{2}}C_\kappa(WU)(dW)$$
$$=\frac{m^{\frac{np}{2}+k}\left(\frac{n}{2}\right)_\kappa\Gamma_p\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}-k\right)}{\Gamma\left(\frac{m}{2}+\frac{np}{2}\right)}C_\kappa(U). \qquad (2.16)$$

Proof. In this case take $h(y)=(1+m^{-1}y)^{-(m+np)/2}$, then

$$S=\frac{m^{\frac{np}{2}+k}\Gamma(\frac{np}{2}+k)\Gamma\left(\frac{m}{2}-k\right)}{\Gamma\left(\frac{m}{2}+\frac{np}{2}\right)}.$$

■

However, we must say that the corresponding value of this integral in Xu and Fang (1989), example 3.2, p.474-476 is incorrect, and in consequence the results (3.1) on expectations given in p.476. The respective general expectation problem involving invariant polynomials of Davis (1980) was proposed without proof by Runze (1997) at the end of page 69, then by a similar procedure of the page 70 we could find the right expression of the cited eq. (3.1) of Xu and Fang (1989); but, again, we find incorrect the last expression of Runze (1997), p.69, and its respective source, the theorem 2.3 part b of Runze (1997). As we shall see, we just need to fill the omitted details in Runze (1997) for checking the errors.

Consider the density of $B=Y'Y$ given in eq. (1.5) of Runze (1997):

$$\frac{\pi^{np/2}c_{n,p}}{\Gamma_p\left(\frac{n}{2}\right)}|\Sigma|^{-n/2}|B|^{(n-p-1)/2}h(\operatorname{tr}\Sigma^{-1}B). \qquad (2.17)$$

where $c_{n,p}$ is a normalization constant (see Runze (1997)). Then another consequence of theorem 2.1.1 is the expectation of a zonal polynomial respect to B with the above defined density.

Corollary 2.1.5. *Suppose that B has density (2.17) and A is an arbitrary symmetric $p \times p$ constant matrix, then*

$$E_B[C_\kappa(BA)] = \frac{\pi^{np/2} c_{n,p} \left(\frac{n}{2}\right)_\kappa}{\Gamma\left(\frac{np}{2}+k\right)} C_\kappa(A\Sigma) I, \tag{2.18}$$

where

$$I = \int_0^\infty h(y) y^{\frac{np}{2}+k-1} dy < \infty. \tag{2.19}$$

Proof. Multiply the density function (2.17) of B by $C_\kappa(BA)$, then integrate over $B > 0$ by using theorem 2.1.1, with $Z = \Sigma^{-1}$, $a = \frac{n}{2}$, $Y = A$ and $m = p$. ∎

The above result is the parallel expression of corollary 2 of Runze (1997), p.68.

Now, let A be an $m \times m$ symmetric matrix with latent roots $a_1, \ldots, a_m$ and recall that zonal polynomials can be expressed in terms of at least five bases (see for example James (1964)).

In particular, recall that the j-th elementary symmetric function r_j of $a_1, \ldots, a_m$ is defined by

$$r_1(A) = \sum_i^m a_i = \operatorname{tr} A,\ r_2(A) = \sum_{i<j}^m a_i a_j, \ldots, r_m(A) = a_1 \cdots a_m = |A|,$$

then zonal polynomials can be written in terms of the r's as follows (see Littlewood (1950))

$$C_\kappa(A) \quad = \quad d_\kappa r_1^{k_1-k_2} r_2^{k_2-k_3} \cdots r_m^{k_m} + \text{lower terms}, \tag{2.20}$$

where "lower terms" is meant monomials in the r_p similar to the one displayed but corresponding to partitions $\tau < \kappa = (k_1, \ldots, k_m)$.

Now, if $\kappa = (1, \ldots, 1)$ is a partition of k and $l_1, \ldots, l_m$ denote the latent roots of A, then

$$\begin{aligned} C_\kappa(A) &= d_\kappa l_1 \cdots l_k + \text{terms of lower weight} \\ &= d_\kappa r_k(A), \end{aligned} \tag{2.21}$$

where $r_k(A)$ is the k-th elementary symmetric function of $l_1, \ldots, l_m$; see Muirhead (1982), p.251.

So we have

Corollary 2.1.6. *Under the conditions of corollary 2.1.5*

$$E_B[r_j(B)] = \frac{\pi^{np/2} c_{n,p} \left(\frac{n}{2}\right)_\kappa}{\Gamma\left(\frac{np}{2} + k\right)} r_j(\Sigma) I, \tag{2.22}$$

where $j = 1, \ldots, p$.

At this point we have studied integrals involving zonal polynomials (with one matrix argument), but recall that the noncentral multivariate analysis was revolutionized by Davis (1980) when extended those polynomials to invariant polynomials of several matrix argument; then a number of distributional results in every field of statistics could be solved.

For example, the expectation under the above defined B of invariant polynomials can be stated; this result was proposed without details by Runze (1997), theorem 2.3, part b; but as we shall see some modifications must be done, for that reason we present the complete proof here as a consequence of our theorem 2.1.1 and corollary 2.1.5.

Corollary 2.1.7. *Suppose that B has density (2.17) and A and R are arbitrary symmetric $p \times p$ constant matrices, then*

$$E_B[C_\phi^{\kappa,\lambda}(BR, BU)] = \frac{\pi^{np/2} c_{n,p} \left(\frac{n}{2}\right)_\phi}{\Gamma\left(\frac{np}{2} + k + l\right)} C_\phi^{\kappa,\lambda}(R\Sigma, U\Sigma) I, \tag{2.23}$$

where

$$I = \int_0^\infty h(y) y^{\frac{np}{2}+k+l-1} dy < \infty, \tag{2.24}$$

κ, λ and ϕ are partitions of k, l and $k+l$, respectively; see Davis (1980) for the theory of invariant polynomials.

Proof. Let $B = \Sigma^{1/2} X \Sigma^{1/2}$, where $(dB) = |\Sigma|^{(p+1)/2}(dU)$; then from (2.17)

$$E_B[C_\phi^{\kappa,\lambda}(BR, BU)] = \frac{\pi^{np/2} c_{n,p}}{\Gamma_p\left(\frac{n}{2}\right)} |\Sigma|^{-n/2}$$
$$\int_{B>0} |B|^{(n-p-1)/2} h(\operatorname{tr} \Sigma^{-1} B) C_\phi^{\kappa,\lambda}(BR, BU)(dB)$$
$$= \frac{\pi^{np/2} c_{n,p}}{\Gamma_p\left(\frac{n}{2}\right)}$$
$$\int_{X>0} |X|^{(n-p-1)/2} h(\operatorname{tr} X) C_\phi^{\kappa,\lambda}(\Sigma^{1/2} X \Sigma^{1/2} R, \Sigma^{1/2} X \Sigma^{1/2} U)(dX)$$
$$= E_X[C_\phi^{\kappa,\lambda}(\Sigma^{1/2} X \Sigma^{1/2} R, \Sigma^{1/2} X \Sigma^{1/2} U)].$$

Note that the distribution of X is invariant under the transformation $X \to H'XH$, $H \in O(p)$, then by Davis (1980) we have that

$$E_X[C_\phi^{\kappa,\lambda}(\Sigma^{1/2} X \Sigma^{1/2} R, \Sigma^{1/2} X \Sigma^{1/2} U)]$$
$$= E_X\left[\int_{O(p)} C_\phi^{\kappa,\lambda}(XH\Sigma^{1/2} R \Sigma^{1/2} H', XH\Sigma^{1/2} U \Sigma^{1/2} H')(dH)\right]$$
$$= \frac{C_\phi^{\kappa,\lambda}(\Sigma^{1/2} R \Sigma^{1/2}, \Sigma^{1/2} U \Sigma^{1/2})}{C_\phi(I_p)} E_X\left[C_\phi(X)\right];$$

and by (2.18)

$$E_X[C_\phi(X)] = \frac{\pi^{np/2} c_{n,p} \left(\frac{n}{2}\right)_\phi}{\Gamma\left(\frac{np}{2}+k+l\right)} C_\phi(I_p) \int_0^\infty h(y) y^{\frac{np}{2}+k+l-1} dy,$$

which establishes the required result. ∎

This is the same result proposed by Runze (1997) except for the uncorrect reference for ϕ in the gamma argument.

As an example of expectations for zonal and invariant polynomials we return to example 3.2 of Xu and Fang (1989), p.474-476 (explicitly their eq. (3.1)) by deriving the correct result by two methods.

First, if we take $h(B) = (1+m^{-1}\operatorname{tr} B)^{-(m+np)/2}$ in corollary 2.1.5 we have that $I = m^{k+np/2}\mathrm{B}(k+np/2, -k+m/2)$, $c_{n,p} = \Gamma[(m+np)/2]/[\pi^{np/2}m^{np/2}\Gamma(m/2)]$ and

$$E_B[C_\kappa(BA)] \quad = \quad \frac{m^k\Gamma[\frac{m}{2}-k]\left(\frac{n}{2}\right)_\kappa}{\Gamma\left(\frac{m}{2}\right)}C_\kappa(A\Sigma). \tag{2.25}$$

For the correctness of the results, note that if we change a little the procedure here derived we obtain the expression for $EC_\kappa(B^{-1}A)$ given by Runze (1997), p. 69.

And second, the corollary 2.1.7 in this case turns

$$E_B[C_\phi^{\kappa,\lambda}(BR, BU)] \quad = \quad \frac{m^{k+l}\Gamma[\frac{m}{2}-k-l]\left(\frac{n}{2}\right)_\phi}{\Gamma(m/2)}C_\phi^{\kappa,\lambda}(R\Sigma, U\Sigma), \tag{2.26}$$

this expression is also computed by Runze (1997), p. 69. however that result is incorrect.

Finally, we can check again (2.25) by using properties of invariant polynomials in (2.26). In fact, just take $R = 0$, then $C_\phi^{\kappa,\lambda}(0, BU) = 0$ for $k > 0$ and $C_\phi^{\kappa,\lambda}(0, BU) = C_\lambda(BU)$ for $k = 0$, then

$$\begin{aligned} E_B[C_\lambda(BU)] &= E_B[C_\phi^{\kappa,\lambda}(0, BU)] \\ &= \frac{m^l\Gamma[\frac{m}{2}-l]\left(\frac{n}{2}\right)_\lambda}{\Gamma(m/2)}C_\lambda(U\Sigma), \end{aligned}$$

which corresponds with (2.25).

Finally, recall that if we take $m = 1$ in the family of Pearson VII distributions we obtain the multivariate Cauchy distribution; thus by replacing this parameter in (2.25) we must have that $\frac{1}{2} \geq k$, which means that the Cauchy distribution has no moments.

2.2 Configuration Density

First, we recall a definition given by Goodall and Mardia (1993) (see section 1.1).

Two figures $X : N \times K$ and $X_1 : N \times K$ have the *same configuration*, or *affine shape*, if $X_1 = XE + 1_N e'$, for some translation $e : K \times 1$ and a nonsingular $E : K \times K$.

The configuration coordinates are constructed in two steps summarized in the expression

$$LX = Y = UE. \tag{2.27}$$

The matrix $U : N-1 \times K$ contains configuration coordinates of X. Let $Y_1 : K \times K$ be nonsingular and $Y_2 : q = N - K - 1 \geq 1 \times K$, such that $Y = (Y_1' \mid Y_2')'$. Define also $U = (I \mid V')'$, then $V = Y_2 Y_1^{-1}$ and $E = Y_1$.

Where L is an $N-1 \times N$ Helmert sub-matrix, see (1.3).

Now we establish an important jacobian:

Lemma 2.2.1. *Let* $(F^{1/2})^2 = F > 0$, $H \in O(K)$, *and* $E = F^{1/2}H$ *so for* $Y = UF^{1/2}H$ *then*

$$(dY) = 2^{-K}|F|^{(q-1)/2}(dV)(dF)(H'dH). \tag{2.28}$$

Proof. Let $E = F^{1/2}H$, with E a $K \times K$ invertible matrix, H orthogonal and $F^{1/2} > 0$. So $E'E = H'FH$, because $E'E$ and F are symmetric and H non singular, then $(E'E) = |H|^{K+1}(dF) = (dF)$. But by theorem 2.1.14 of Muirhead (1982): $(dE) = 2^{-K}|E'E|^{-1/2}(E'E)(H'dH)$. Then we obtain: $(dE) = 2^{-K}|H'FH|^{-1/2}(dF)(H'dH) = 2^{-K}|F|^{-1/2}(dF)(H'dH)$. Summarizing we get

$$E = F^{1/2}H \Rightarrow (dE) = 2^{-K}|F|^{-1/2}(dF)(H'dH). \tag{2.29}$$

Now,

$$Y = \begin{pmatrix} I \\ V \end{pmatrix} E = \begin{pmatrix} E \\ VE \end{pmatrix}.$$

Differentiating and computing the exterior product, we get
$(dY) = |E|^q(dV)(dE)$, but $|E| = |F^{1/2}H| = |F|^{1/2}$, so

$$(dY) = |F|^{q/2}(dV)(dE). \tag{2.30}$$

Replacing (2.29) in (2.30) we obtain the required result. ■

Now we can state with the help of theorem 2.1.1 the main statistical result of this work, the general case of the configuration density under a non-isotropic noncentral elliptical model.

Theorem 2.2.1. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1} \otimes I_K, h)$, for Σ positive definite ($\Sigma > 0$), $\mu \neq 0_{N-1\times K}$, then the configuration density is given by*

$$\frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \sum_{t=0}^{\infty} \frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)} \sum_{r=0}^{\infty} \frac{1}{r!}\left[\mathrm{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r$$

$$\sum_{\tau} \frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau} C_\tau(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1})S. \tag{2.31}$$

where

$$S = \int_0^\infty h^{(2t+r)}(y) y^{\frac{K(N-1)}{2}+t-1} dy < \infty, \tag{2.32}$$

Proof. The density of Y is given by

$$\frac{1}{|\Sigma|^{\frac{K}{2}}}\, h[\mathrm{tr}[(Y-\mu)'\Sigma^{-1}(Y-\mu)]], \tag{2.33}$$

If we factorize Y according to lemma 2.2.1, then the joint density of U, F and H is

$$\frac{2^{-K}|F|^{(q-1)/2}}{|\Sigma|^{\frac{K}{2}}} \tag{2.34}$$
$$h\left[\mathrm{tr}\left(\mu'\Sigma^{-1}\mu + FU'\Sigma^{-1}U\right) + \mathrm{tr}\left(-2\mu'\Sigma^{-1}UF^{1/2}H\right)\right](H'dH)(dF)(dV).$$

Assuming that h admits a Taylor expansion (see Fang (1990a,b)), the joint density of U, F and H becomes:

$$\frac{2^{-K}|F|^{(q-1)/2}}{|\Sigma|^{\frac{K}{2}}}\sum_{t=0}^{\infty}\frac{1}{t!}h^{(t)}\left(\mathrm{tr}\left(FU'\Sigma^{-1}U\right) + \mathrm{tr}\left(\mu'\Sigma^{-1}\mu\right)\right)$$
$$\left[\mathrm{tr}\left(-2\mu'\Sigma^{-1}UF^{1/2}H\right)\right]^t (H'dH)(dF)(dV).$$

Now, recall that

$$\int_{O(m)}[\mathrm{tr}(XH)]^r(dH) = 0 \quad \text{for odd } r, \tag{2.35}$$

and

$$\int_{O(m)}[\mathrm{tr}(XH)]^{2k}(dH) = \sum_{\kappa}\frac{\left(\frac{1}{2}\right)_k}{\left(\frac{1}{2}m\right)_\kappa}C_\kappa(XX'), \tag{2.36}$$

(see James (1964)) then integration with respect to H gives the joint density of F and U as follows

$$\frac{\pi^{K^2/2}|F|^{(q-1)/2}}{|\Sigma|^{\frac{K}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!}h^{(2t)}\left(\mathrm{tr}\left(FU'\Sigma^{-1}U\right) + \mathrm{tr}\left(\mu'\Sigma^{-1}\mu\right)\right)$$
$$\sum_{\tau}\frac{1}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}UF\right)(dF)(dV). \tag{2.37}$$

And noting that $h^{2t}(\cdot)$ admits a Taylor expansion, then the joint density of F and U finally takes the form:

$$\frac{\pi^{K^2/2}|F|^{(q-1)/2}}{|\Sigma|^{\frac{K}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}h^{(2t+r)}\left(\operatorname{tr}\left(FU'\Sigma^{-1}U\right)\right)\left[\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r$$
$$\sum_{\tau}\frac{1}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}UF\right)(dF)(dV).$$

So integration over $F>0$ gives the configuration density as follows

$$\frac{\pi^{K^2/2}}{|\Sigma|^{\frac{K}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!}\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r\sum_{\tau}\frac{1}{\left(\frac{K}{2}\right)_\tau} \tag{2.38}$$
$$\int_{F>0}h^{(2t+r)}\left(\operatorname{tr}\left(FU'\Sigma^{-1}U\right)\right)|F|^{(q-1)/2}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}UF\right)(dF).$$

Goodall and Mardia (1993), corrected by Díaz-García *et al.* (2003), derived the configuration density in the normal case, by evaluating the above multivariate integral using the classical results of Constantine (1963) (in Muirhead (1982)). So, it is clear that the advances in this kind of problems depend on evaluation of this multiple integral for non gaussian cases, and only a few of them are available.

However, if we compare (2.38) and (2.4) we find that the configuration density of any elliptical model reduces to the computation of a single integral, simplifying the above multivariate problem into a simple ordinary integration (providing that $h^{(2t+r)}(y)$ and the improper integral exist, of course).

So, the integral in (2.38) is reduced by (2.4) to

$$\int_{F>0}h^{(2t+r)}\left(\operatorname{tr}\left(FU'\Sigma^{-1}U\right)\right)|F|^{(q-1)/2}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}UF\right)(dF)$$
$$=\frac{|U'\Sigma^{-1}U|^{-\frac{N-1}{2}}\left(\frac{N-1}{2}\right)_\tau\Gamma_K\left(\frac{N-1}{2}\right)C_\tau(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1})}{\Gamma\left(\frac{K(N-1)}{2}+t\right)}S,$$

where

$$S=\int_0^\infty h^{(2t+r)}(y)y^{\frac{K(N-1)}{2}+t-1}dy<\infty,$$

and the required result follows. ∎

The central non-isotropic configuration density must receive an special attention (see Díaz-García *et al.* (2003) for the central isotropic case).

First we need a preliminary result which ratifies our main assertion in this work, the multiple elliptical integrals (if exist) over positive definite matrix X with kernel $f(\operatorname{tr} ZX)$ might reduce to the computation of a single integral (if exist).

Lemma 2.2.2. *Let Z be a complex symmetric $m \times m$ matrix with $Re(Z) > 0$. Then*

$$\begin{aligned}\int_{X>0} \quad & h(\operatorname{tr} XZ)|X|^{a-(m+1)/2}(dX) \\ & = \frac{|Z|^{-a}\Gamma_m(a)}{\Gamma(ma)} S, \end{aligned} \tag{2.39}$$

where

$$S = \int_0^\infty h(y) y^{ma-1} dy < \infty. \tag{2.40}$$

Proof. We give three proofs of this result. A first heuristic proof comes from considering (2.4). For $k > 0$ and $Y = 0$, $C_\kappa YX = 0$, so there is nothing to prove. But if $k = 0$ and $Y = 0$, then by definition $(a)_\kappa = 1$, and $C_\kappa(A) = 1$ (both heuristic-trivial definitions can be seen equivalent to the heuristic definition $0! = 1$, see Muirhead (1982) p.258, and Xu and Fang (1989), p.477), the result follows.

A second heuristic proof comes from eq. (2.5) of Runze (1997), where $X = W$, $Y = U$, $Z = I_p$, apply the same definitions recalling that $\Gamma_p(a) = \pi^{p(p-1)/4} \prod_{i=1}^p \Gamma[a - (i-1)/2]$ and the same result is found.

In a third proof we just solve the multiple integral without zonal polynomials. As it was establish historically (Wishart) the multivariate gamma function around the $20's$ (preceding the zonal's integrals by the 60's). The "normalization" procedure is typical, suppose that $Z > 0$ is real. Put $X = Z^{-1/2}VZ^{-1/2}$, so that $(dX) = |Z|^{-(m+1)/2}(dV)$ and the l.h.s of (2.39) becomes

$$\begin{aligned} &\int_{X>0} h(\operatorname{tr} XZ)|X|^{a-(m+1)/2}(dX) \\ = \quad & |Z|^{-a}\int_{V>0} h(\operatorname{tr} V)|V|^{a-(m+1)/2}(dV). \end{aligned} \tag{2.41}$$

To solve the above integral we proceed as in proof of (2.4). put $V = T'T$, where T is upper triangular with positive diagonal elements and
$(dV) = 2^m \prod_{i=1}^m t_{ii}^{m+1-i}(dT)$; then $\operatorname{tr} V = \sum_{i\leq j}^m t_{ij}^2$, $|V| = \prod_{i=1}^m t_{ii}^2$.
Hence by Caro *and* Nagar (2006) or Fang and Zhang (1990) we have, after some simplification, that

$$\begin{aligned} \int_{V>0} h(\operatorname{tr} V)|V|^{a-(m+1)/2}(dV) &= \int\cdots\int_{t_{ij}} h\left(\sum_{i\leq j}^m t_{ij}^2\right)\prod_{i=1}^m t_{ii}^{2a-i}2^m(dT) \\ &= \frac{\prod_{i=1}^m \Gamma\left(a-\frac{1}{2}(i-1)\right)\pi^{\frac{1}{4}m(m-1)}}{\Gamma(ma)}\int_0^\infty h(y)y^{ma-1}dy, \end{aligned}$$

and using this in (2.41), we get the required result for real Z; and it follows for complex Z by analytic continuation. Since $\operatorname{Re}(Z) > 0$, $|Z| \neq 0$ and $|Z|^a$ is well defined in continuation. ■

Note that if we take $Z = \Sigma^{-1}$ and $h(y) = e^{-y/2}$ which implies $S = 2^{ma}\Gamma(ma)$; we have the classical result for multivariate gamma function proved in detail by Muirhead (1982), p.61-63.

$$\begin{aligned} &\int_{A>0} \operatorname{etr}(-\frac{1}{2}\Sigma^{-1}A)|A|^{a-(m+1)/2}(dA) \\ &\quad = \prod_{i=1}^m \Gamma\left(a-\frac{1}{2}(i-1)\right)\pi^{\frac{1}{4}m(m-1)}|\Sigma|^a 2^{ma}. \end{aligned} \tag{2.42}$$

So, finally the central case of the configuration density follows easily from lemma 2.2.2.

Corollary 2.2.1. *If* $Y \sim E_{N-1\times K}(0_{N-1\times K}, \Sigma_{N-1\times N-1} \otimes I_K, h)$*, for* $\Sigma > 0$*, then the central configuration density is invariant under the elliptically contoured distributions and it is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{\frac{Kq}{2}}|\Sigma|^{\frac{K}{2}}\Gamma_K\left(\frac{K}{2}\right)}|U'\Sigma^{-1}U|^{-\frac{N-1}{2}}. \tag{2.43}$$

The preceding proposition generalizes theorem 3.2 of Díaz-García *et al.* (2003) (see (1.5)) which concerned the isotropic case, $\Sigma = \sigma^2 I_{N-1}$, (recall that $|U'U| = |I_K + V'V|$); in this case both expression coincide excepting their factor 2^K, since Goodall and Mardia (1993) and Díaz-García *et al.* (2003) did not described the Haar measure $(H'dH)$ employed in the computation of the jacobian of $Y = UF^{1/2}H$, see our lemma 2.2.1 for solving this discrepancy.

Most of the applications in statistical theory of shape resides on the isotropic model (see Dryden and Mardia (1998)), so in the case of the noncentral elliptical configuration density if we take $\Sigma = \sigma^2 I_{N-1}$ in theorem 2.2.1 we obtain

Corollary 2.2.2. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$*, then the isotropic noncentral configuration density is given by*

$$\frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)}\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)\right]^r$$
$$\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right)S, \tag{2.44}$$

where

$$S = \int_0^\infty h^{(2t+r)}(y)y^{\frac{K(N-1)}{2}+t-1}dy < \infty. \tag{2.45}$$

Remark 2.2.1. We have seen that (2.4) and Runze (1997), eq. (2.5) provides parallel results very closed related. So we could try defining a new configuration density based on another factorization which lead to those integrals of Runze and to replicate our procedure for finding the new distribution. But unfortunately, the conditions on $q = N - K - 1 \geq 1$ and the conditions for gamma's provides useless integrals for configurations. Explicitly, if we factorized Y by $Y = UF^{-1/2}H$ so that $(dY) = 2^{-K}|F|^{-(q-1)-(K+1)}(dV)(H'dH)(dF)$ which supposes absurd constraint $q \leq -K$ in the shape theory context. By the other hand, note that distributions type eq. (2.5) of Runze (1997) are valid only where $a > k_1 + (p-1)/2$ important restriction when Taylor series expansions are considered.

Remark 2.2.2. A new enrichment family of shape distributions can be proposed if we replace (2.4) by the following integral. Let Y be a symmetric $m \times m$ matrix, $\mathrm{Re}(a) > (m-1)/2$, $\mathrm{Re}(b) > (m-1)/2$

$$\begin{aligned} & \int_{0<Z<I_m} |Z|^{a-\frac{1}{2}(m+1)}|I-Z|^{b-\frac{1}{2}(m+1)} \\ & \int_{X>0} h(\operatorname{tr} XZ)h(\operatorname{tr} X(I-Z))|X|^{a+b-(m+1)/2}(dX)C_\kappa(ZY)(dZ) \\ = & \ \frac{C_\kappa(Y)}{C_\kappa(I_m)}S, \end{aligned}$$

where

$$\begin{aligned} S = \int_{0<Z<I_m}\int_{X>0} & |Z|^{a-\frac{1}{2}(m+1)}|I-Z|^{b-\frac{1}{2}(m+1)}|X|^{a+b-(m+1)/2} \\ & h(\operatorname{tr} XZ)h(\operatorname{tr} X(I-Z))C_\kappa(Z)(dZ)(dX), \end{aligned}$$

and the respective simplification of S with the general technique studied for (2.4). Then multivariate Beta type integrals can be included in the repertory of shape distributions. However we leave this study for a subsequent work.

Chapter 3

Families of elliptical configuration densities

We already found the noncentral non-isotropic configuration density with a remarkable property, given an elliptically contoured distribution, we do not need to compute as usual a multivariate integral but a single integral. However, the difficulty sometimes will arise on the derivatives and the series simplifications involved. But as we shall see all the classical elliptical families can be computed explicitly and all of them will be suitable for inference on the exact distribution.

In this chapter we derive explicit configuration densities for matrix variate symmetric Kotz type distributions (it includes normal), matrix variate Pearson type VII distributions (it includes t and Cauchy distributions), matrix variate symmetric Bessel distribution (it includes Laplace distribution) and matrix variate symmetric Jensen-logistic distribution.

We must refer for applications the posterior study of the isotropic case because it is the more widely used in all the applications of shape theory, see Dryden and Mardia (1998).

3.1 Normal configuration density

We start with this case because it is the simplest one and the only noncentral (isotropic) published result. Goodall and Mardia (1993) proposed that density but it was corrected by Díaz-García *et al.* (2003), however an imprecision in the information of the haar measure employed in both derivations motivated our lemma 2.2.1 and we have to clarify this distribution again. As we shall see it will be derived by three different methods, as direct corollary of theorem 2.2.1 and as two consequences of the Kotz distribution.

Recall that the $p \times n$ random matrix X is said to have a matrix variate normal distribution with parameters, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}|\Phi|^{p/2}} \operatorname{etr}\left[-\frac{1}{2}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right].$$

The first method is given below.

Corollary 3.1.1. *If $Y \sim \mathcal{N}_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1} \otimes I_K)$, for $\Sigma > 0$, then the non-isotropic noncentral normal configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right) \operatorname{etr}\left(\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1} - \frac{1}{2}\mu'\Sigma^{-1}\mu\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$${}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right). \tag{3.1}$$

Proof. See (2.31), so for the normal case take

$$h(y) = \frac{1}{(2\pi)^{K(N-1)/2}} \mathrm{e}^{-\frac{1}{2}y}$$

thus

$$h(y)^{(2t+r)} = \frac{(-2)^{-2t-r}}{(2\pi)^{K(N-1)/2}} \mathrm{e}^{-\frac{1}{2}y},$$

and (2.32) becomes

$$S = \frac{\Gamma\left(\frac{K(N-1)}{2} + t\right)}{(-1)^r 2^{t+r} \pi^{K(N-1)/2}}.$$

Replacing S in (2.31), using $aC_\kappa(A) = C_\kappa(a^k A)$ and $q = N - 1 - K \geq 1$, we have

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \sum_{t=0}^{\infty} \frac{1}{t!} \left\{ \sum_{r=0}^{\infty} \frac{1}{r!} \left[\mathrm{tr}\left(-\frac{1}{2}\mu'\Sigma^{-1}\mu\right)\right]^r \right\}$$
$$\sum_\tau \frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau} C_\tau\left(\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right).$$

The term in braces is clearly $\exp(\cdot)$ which does not depend on t and if we use the definition of hypergeometric function of matrix arguments (see (1.20)), explicitly ${}_1F_1(a; b; X) = \sum_{t=0}^{\infty} \frac{1}{t!} \sum_\tau \frac{(a)_\tau}{(b)_\tau} C_\tau(X)$, we have

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right) \mathrm{etr}\left(-\frac{1}{2}\mu'\Sigma^{-1}\mu\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$${}_1F_1\left(\frac{N-1}{2}; \frac{K}{2}; \frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right).$$

The normal non-isotropic configuration density follows by the Kummer relations

$${}_1F_1(a; c; X) = \mathrm{etr}\, X {}_1F_1(c - a; c; -X), \tag{3.2}$$

see (1.28), i.e.

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right) \mathrm{etr}\left(\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1} - \frac{1}{2}\mu'\Sigma^{-1}\mu\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$${}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right).$$

■

Now, if we take $\Sigma = \sigma^2 I_{N-1}$, the isotropic case, we get

Corollary 3.1.2. *If $Y \sim \mathcal{N}_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K)$, then the isotropic noncentral normal configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K+V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \operatorname{etr}\left(\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu - \frac{1}{2\sigma^2}\mu'\mu\right)$$
$$ {}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu\right). \tag{3.3}$$

See (1.4), this is the same result of Díaz-García *et al.* (2003) (proposed by Goodall and Mardia (1993) with some errors) except for the factor 2^k which comes from their anonymous jacobian computation (see (2.28)), both woks computed the configuration density by using classical results of Muirhead (1982) (Constantine (1963)) which supposes integration over $F > 0$, note that we replace that by using theorem 2.1.1 which requires the computation of a single integral, in this case the matrix S in (2.32). Of course, the multiple integrals used by the cited works are consequence of theorem 2.1.1 as we proved in corollary 2.1.2 and 2.1.3.

We shall derive (3.1) again as a consequence of the Kotz non-isotropic configuration density.

3.2 Pearson type VII configuration density

In this section we derive the non-isotropic noncentral Pearson type VII configuration density as a simple consequence of (2.31) instead of performing the classical multiple integration, suggested by this kind of distribution.

Recall that a $p\times n$ random matrix X is said to have a matrix variate symmetric Pearson type VII distribution with parameters $s, R \in \mathbb{R}$, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $R > 0$, $s > np/2$, $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{\Gamma(s)}{(\pi R)^{np/2}\Gamma\left(s-\frac{np}{2}\right)|\Sigma|^{n/2}|\Phi|^{p/2}}\left(1+\frac{\operatorname{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}}{R}\right)^{-s}.$$

When $s=(np+R)/2$, X is said to have a matrix variate t-distribution with R degrees of freedom. And in this case, if $R=1$, then X is said to have a matrix variate Cauchy distribution.

We already discussed some facts of this distributions in subsection 2.1.1, see corollary 2.1.4, and (2.25).

Now the corresponding configuration is derived.

Corollary 3.2.1. *If* $Y\sim E_{N-1\times K}(\mu_{N-1\times K},\Sigma_{N-1\times N-1}\otimes I_K,h)$, *for* $\Sigma>0$, *then the non-isotropic noncentral Pearson type VII configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$${}_1P_1\left(\left(s-\frac{K(N-1)}{2}\right)_t\left(1+\frac{\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)}{R}\right)^{-s+\frac{K(N-1)}{2}-t}:\right.$$
$$\left.\frac{N-1}{2};\frac{K}{2};\frac{1}{R}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right), \tag{3.4}$$

where

$${}_1P_1(f(t):a;b;X)=\sum_{t=0}^{\infty}\frac{f(t)}{t!}\sum_{\tau}\frac{(a)_\tau}{(b)_\tau}C_\tau(X), \tag{3.5}$$

Proof. In this case we take

$$h(y)=\frac{\Gamma(s)}{(\pi R)^{\frac{K(N-1)}{2}}\Gamma\left(s-\frac{K(N-1)}{2}\right)}\left(1+\frac{y}{R}\right)^{-s},$$

then

$$h(y)^{(2t+r)}=\frac{\Gamma(s)(-1)^r(s)_{2t+r}}{(\pi R)^{\frac{K(N-1)}{2}}\Gamma\left(s-\frac{K(N-1)}{2}\right)R^{2t+r}}\left(1+\frac{y}{R}\right)^{-(s+2t+r)},$$

and after some simplification, (2.32) becomes

$$\mathcal{S} = \frac{(-1)^r \Gamma\left(\frac{K(N-1)}{2} + t\right)\left(s - \frac{K(N-1)}{2}\right)_{t+r}}{\pi^{\frac{K(N-1)}{2}} R^{t+r}}.$$

Thus (2.31) takes the form

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$\sum_{t=0}^{\infty}\frac{1}{t!}\left\{\sum_{r=0}^{\infty}\frac{\left(s-\frac{K(N-1)}{2}\right)_{t+r}}{r!}\left[\mathrm{tr}\left(-\frac{1}{R}\mu'\Sigma^{-1}\mu\right)\right]^r\right\}$$
$$\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{R}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right).$$

Surprisingly, as in the normal case the term in braces preserves the core of the distribution (this does not happen in the remaining configurations), in this case it is $(a)_t(1+u/R)^{-a-t}$, with $a = s - K(N-1)/2 > 0$ (see the above definition of the Pearson type VII distribution), $u = \mathrm{tr}(\mu'\Sigma^{-1}\mu)$, so we get the non-isotropic noncentral Pearson type VII configuration density as follows

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$\sum_{t=0}^{\infty}\frac{\left(s-\frac{K(N-1)}{2}\right)_t\left(1+\frac{\mathrm{tr}(\mu'\Sigma^{-1}\mu)}{R}\right)^{-s+\frac{K(N-1)}{2}-t}}{t!}$$
$$\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{R}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right).$$

■

Remark 3.2.1. The similarity of the preceding series with the hypergeometric functions on the normal case motivates us the following notation

$$_1P_1(f(t):a;b;X)=\sum_{t=0}^{\infty}\frac{f(t)}{t!}\sum_{\tau}\frac{(a)_\tau}{(b)_\tau}C_\tau(X),$$

where

$$_1P_1(f(t):a;b;X)\;=\;{}_1F_1(1:a;b;X).$$

Given the above generalization of the hypergeometric function it should be important a further study of the series $_1P_1(f(t):a;b;X)$, and we leave it for a subsequent work.

If we take $s=(K(N-1)+R)/2$ in corollary 3.2.1 we obtain the configuration density associated to a matrix variate t-distribution with R degrees of freedom.

Corollary 3.2.2. *If* $Y\sim E_{N-1\times K}(\mu_{N-1\times K},\Sigma_{N-1\times N-1}\otimes I_K,h)$, *with* $\Sigma>0$, *then the non-isotropic noncentral* t *configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$_1P_1\left(\left(\frac{R}{2}\right)_t\left(1+\frac{\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)}{R}\right)^{-\frac{R}{2}-t}:\right.$$
$$\left.\frac{N-1}{2};\frac{K}{2};\frac{1}{R}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right). \qquad (3.6)$$

And if we replace $R=1$ in the above density, we obtain the respective Cauchy configuration density.

Corollary 3.2.3. *If* $Y\sim E_{N-1\times K}(\mu_{N-1\times K},\Sigma_{N-1\times N-1}\otimes I_K,h)$, *with* $\Sigma>0$, *then the non-isotropic noncentral Cauchy configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$_1P_1\left(\left(\frac{1}{2}\right)_t\left(1+\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right)^{-\frac{1}{2}-t}:\right.$$
$$\left.\frac{N-1}{2};\frac{K}{2};U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right). \tag{3.7}$$

3.3 Kotz type configuration density

In this section we study in detail the full non-isotropic noncentral Kotz configuration density by using some results in partition theory, which are the key for applying (2.31) for all the parameters of the family.

First, the class of elliptically contoured distribution is defined. The $p\times n$ random matrix X is said to have a matrix variate symmetric Kotz type distribution with parameters $T, R, s \in \Re$, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $R > 0$, $s > 0$ $2T + np > 2$, $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{sR^{\frac{2T+np-2}{2s}}\Gamma\left(\frac{np}{2}\right)}{\pi^{np/2}\Gamma\left(\frac{2T+np-2}{2s}\right)|\Sigma|^{n/2}|\Phi|^{p/2}}\left[\operatorname{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right]^{T-1}$$
$$\exp\left\{-R\operatorname{tr}\left[(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right]^s\right\}. \tag{3.8}$$

When $T = s = 1$, and $R = 1/2$ we get the probability density function of the absolutely continuous matrix variate normal distribution.

Now, for finding a closed form of (2.31) we need some additional theory, in this case the partition theory provides suitable expressions for derivatives, we use here the results by Caro-Lopera *et al* (2008), chapter 2, see also *Faà di Bruno*'s formula.

Lemma 3.3.1. *Let $h(t) = s(t)f(g(t))$, where s, f and g have derivatives of all orders, if $w^{(k)}$ denotes $\frac{d^k w(t)}{dt^k}$ then*

$$h^{(k)} = \sum_{m=0}^{k} \binom{k}{m} s^{(m)}[f(g(t))]^{(k-m)}, \tag{3.9}$$

where

$$[f(g(t))]^{(k)} = \sum_{\kappa=(k^{\nu_k},(k-1)^{\nu_{k-1}},\cdots,3^{\nu_3},2^{\nu_2},1^{\nu_1})} \frac{k!}{\prod_{i=1}^{k}\nu_i!(i!)^{\nu_i}} f^{(\sum_{i=1}^{k}\nu_i)} \prod_{i=1}^{k}(g^{(i)})^{\nu_i}. \tag{3.10}$$

Note that the functions h considered in the elliptically contoured distributions admits Taylor expansions then the above expressions always exists for all k.

A particular important case of the above lemma applies for the Kotz type distribution $h(y) = cy^{T-1}\,\mathrm{e}^{-Ry^s}$, where c is the normalization constant, $R > 0$, $s > 0$ and $2T + K(N-1) > 2$; (recall that K and N, are the dimension and number of landmarks, as usual in the configuration density and they are related by $q = N - 1 - K \geq 1$).

In order to simplify the expressions due to the preceding lemma, we divide the study of the Kotz type configuration density in three subfamilies:
1. $s = 1, T > 1-K(N-1)/2$; 2. $T = 1, s > 0$; 3. $s \neq 1, T \neq 1$, see Caro-Lopera *et al* (2008), chapter 2. We refer them by Kotz type I, Kotz type II and Kotz type III, respectively. However we can take off the restriction of case 3 and it serves for the general expression, but we consider the division for a didactic reason and for showing easily the main corollary, the normal case.

Lemma 3.3.2. *Let $h(y) = y^{T-1}\,\mathrm{e}^{-Ry^s}$, where, $R > 0$, $s > 0$, $2T + K(N-1) > 2$ and $N - 1 - K \geq 1$; if $w^{(k)}$ denotes $\frac{d^k w}{dy^k}$ and $\sum_{\kappa\in P_r}$ is the summation over all the partitions $\kappa = (k^{\nu_k}, (k-1)^{\nu_{k-1}}, \cdots, 3^{\nu_3}, 2^{\nu_2}, 1^{\nu_1})$ of r, then we have*

- *Case 1:* $s = 1$, $h(y) = y^{T-1}\,\mathrm{e}^{-Ry}$,

$$h^{(k)} = (-R)^k y^{T-1}\,\mathrm{e}^{-Ry}\left\{1 + \sum_{m=1}^{k}\binom{k}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right](-Ry)^{-m}\right\}. \tag{3.11}$$

- *Case 2:* $T = 1$, $h(y) = \mathrm{e}^{-Ry^s}$,

$$h^{(k)} = \mathrm{e}^{-Ry^s}\sum_{\kappa\in P_k}\frac{k!(-R)^{\sum_{i=1}^{k}\nu_i}\prod_{j=0}^{k-1}(s-j)^{\sum_{i=j+1}^{k}\nu_i}}{\prod_{i=1}^{k}\nu_i!(i!)^{\nu_i}}y^{\sum_{i=1}^{k}(s-i)\nu_i}. \tag{3.12}$$

- *Case 3:* $T \neq 1$, $s \neq 1$, $h(y) = y^{T-1}\,\mathrm{e}^{-Ry^s}$,

$$\begin{aligned} h^{(k)} = y^{T-1}\,\mathrm{e}^{-Ry^s}\Bigg\{&\sum_{\kappa\in P_k}\frac{k!(-R)^{\sum_{i=1}^{k}\nu_i}\prod_{j=0}^{k-1}(s-j)^{\sum_{i=j+1}^{k}\nu_i}}{\prod_{i=1}^{k}\nu_i!(i!)^{\nu_i}}y^{\sum_{i=1}^{k}(s-i)\nu_i} \\ &+\sum_{m=1}^{k}\binom{k}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right] \\ &\sum_{\kappa\in P_{k-m}}\frac{(k-m)!(-R)^{\sum_{i=1}^{k-m}\nu_i}\prod_{j=0}^{k-m-1}(s-j)^{\sum_{i=j+1}^{k-m}\nu_i}}{\prod_{i=1}^{k-m}\nu_i!(i!)^{\nu_i}}y^{\sum_{i=1}^{k-m}(s-i)\nu_i-m}\Bigg\}. \end{aligned} \tag{3.13}$$

So with these expressions the Kotz type configuration density can be found in a closed form.

3.3.1 Kotz type I configuration density

In this case the subfamily comes from $s = 1$, then

$$h(y) = \frac{R^{T-1+\frac{K(N-1)}{2}}\Gamma\left(\frac{K(N-1)}{2}\right)}{\pi^{K(N-1)/2}\Gamma\left(T-1+\frac{K(N-1)}{2}\right)}y^{T-1}\,\mathrm{e}^{-Ry},$$

see (3.8), so by using (3.11), after some simplification, and provided that the gamma's exist (by suitable conditions on the corresponding parameters), the integral S in (2.32) becomes

$$S=\frac{(-1)^r R^{t+r}\Gamma\left(\frac{K(N-1)}{2}\right)}{\pi^{K(N-1)/2}\Gamma\left(T-1+\frac{K(N-1)}{2}\right)}\left\{\Gamma\left(T-1+\frac{K(N-1)}{2}+t\right)\right.$$
$$\left.+\sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right](-1)^m\Gamma\left(T-1-m+\frac{K(N-1)}{2}+t\right)\right\}.$$

Thus replacing this expression in (2.31), we have

Corollary 3.3.1. *If $Y\sim E_{N-1\times K}(\mu_{N-1\times K},\Sigma_{N-1\times N-1}\otimes I_K,h)$, with $\Sigma>0$, then the Kotz type I non-isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(T-1+\frac{K(N-1)}{2}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(-R\mu'\Sigma^{-1}\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(RU'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right)$$
$$\left\{\Gamma\left(T-1+\frac{K(N-1)}{2}+t\right)\right. \tag{3.14}$$
$$\left.+\sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right](-1)^m\Gamma\left(T-1-m+\frac{K(N-1)}{2}+t\right)\right\}.$$

Remark 3.3.1. Note that we demand that $T-1-m+\frac{K(N-1)}{2}+t>0$, in other words, in order to perform inference we need to truncate the above series at some t and r, so the Kotz type I configuration density will be defined for $T>1+t+r-\frac{K(N-1)}{2}$ (recall that the elliptical Kotz type I density demands only that $T>1-\frac{K(N-1)}{2}$).

Now, consider $T=1$ in (3.14), then a confluent hypergeometric class of densities indexed by R are obtained, i.e.

Corollary 3.3.2. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1}\otimes I_K, h)$, with $\Sigma > 0$ and $T = 1$, then the Kotz type I non-isotropic noncentral configuration density simplifies to*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)\operatorname{etr}\left(RU'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1} - R\mu'\Sigma^{-1}\mu\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$_1F_1\left(-\frac{q}{2};\frac{K}{2};-RU'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right) \tag{3.15}$$

Proof. When $T = 1$ in (3.14) the summation in the braces over $m = 1,\dots$ clearly vanishes. Then the series in r goes to exponential, independent of t, thus the confluent hypergeometric definition can be applied and the result follows by Kummer relations (3.2). ∎

Finally the normal configuration density can be derived again, but in this case, as a simple consequence of the preceding result, by taking $R = \frac{1}{2}$.

Corollary 3.3.3. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1}\otimes I_K, h)$, with $\Sigma > 0$, $T = 1$, $R = \frac{1}{2}$ then the Kotz type I non-isotropic noncentral configuration density simplifies to normal configuration density and it is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)\operatorname{etr}\left(\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1} - \frac{1}{2}\mu'\Sigma^{-1}\mu\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$_1F_1\left(-\frac{q}{2};\frac{K}{2};-\frac{1}{2}U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right), \tag{3.16}$$

see corollary 3.1.1.

3.3.2 Kotz type II configuration density

The corresponding subfamily comes from $T = 1$ in (3.8), then

$$h(y) = \frac{sR^{\frac{K(N-1)}{2s}}\Gamma\left(\frac{K(N-1)}{2}\right)}{\pi^{K(N-1)/2}\Gamma\left(\frac{K(N-1)}{2s}\right)}\,\mathrm{e}^{-Ry^s},$$

and by using (3.12), the integral S in (2.32) is

$$S = \frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{\pi^{K(N-1)/2}\Gamma\left(\frac{K(N-1)}{2s}\right)} \sum_{\kappa\in P_{2t+r}} \frac{(2t+r)!(-1)^{\sum_{i=1}^{2t+r}\nu_i}\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i}}{\prod_{i=1}^{2t+r}\nu_i!(i!)^{\nu_i}} R^{\frac{\sum_{i=1}^{2t+r} i\nu_i - t}{s}}\Gamma\left(\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i + K(N-1)+2t}{2s}\right); \quad (3.17)$$

and the particular configuration density follows:

Corollary 3.3.4. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1}\otimes I_K, h)$*, with* $\Sigma > 0$*, then the Kotz type II non-isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \sum_{t=0}^{\infty}\frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(\frac{K(N-1)}{2s}\right)} \sum_{r=0}^{\infty}\frac{1}{r!}\left[\mathrm{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r \sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau} C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right) \sum_{\kappa\in P_{2t+r}}\frac{(2t+r)!\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i}\Gamma\left(\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i+K(N-1)+2t}{2s}\right)}{(-1)^{\sum_{i=1}^{2t+r}\nu_i}R^{-\frac{\sum_{i=1}^{2t+r}i\nu_i-t}{s}}\prod_{i=1}^{2t+r}\nu_i!(i!)^{\nu_i}}. \quad (3.18)$$

Again, for the existence of the density we proceed as in remark 3.3.1, by truncating the series at some t and r and then imposing the restriction on s in terms of t, r, K, N, in order to obtain $\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i+K(N-1)+2t}{2s} > 0$

We can get again corollary 3.3.2, and for third time the normal case, by taking $s = 1$ in (3.18) and noting that in this trivial case $\nu_1 = 2t + r$, $\nu_i = 0$ for $i = 2, \ldots, 2t + r$, and $\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i} = (1-0)^{\sum_{i=0+1}^{2t+r}\nu_i} = 1^{2t+r} = 1$. Then the last summation of (3.18) becomes $(-1)^r R^{t+r}\Gamma\left(\frac{K(N-1)}{2}+t\right)$ and the result follows easily by exponential series, hypergeometric function definition and Kummer formula.

3.3.3 Kotz type III configuration density

Finally, the last subfamily which completes the study of the Kotz configuration density can be established with the same technique.

In this case $T \neq 1$ and $s \neq 1$ so from (3.8),

$$h(y) = \frac{sR^{\frac{2T+K(N-1)-2}{2s}}\Gamma\left(\frac{K(N-1)}{2}\right)}{\pi^{K(N-1)/2}\Gamma\left(\frac{2T+K(N-1)-2}{2s}\right)} y^{T-1}\,\mathrm{e}^{-Ry^s},$$

and from (3.13) we have the required derivative; then the integral (2.32) becomes

$$\begin{aligned} S = {} & \frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{\pi^{K(N-1)/2}\Gamma\left(\frac{2T+K(N-1)-2}{2s}\right)} \\ & \left\{\sum_{\kappa\in P_{2t+r}} \frac{(2t+r)!\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i}\Gamma\left(\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i+2T-2+K(N-1)+2t}{2s}\right)}{(-1)^{\sum_{i=1}^{2t+r}\nu_i}R^{\frac{-\sum_{i=1}^{2t+r}i\nu_i+t}{s}}\prod_{i=1}^{2t+r}\nu_i!(i!)^{\nu_i}} \right. \\ & + \sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right] \\ & \sum_{\kappa\in P_{2t+r-m}} \frac{(2t+r-m)!\prod_{j=0}^{2t+r-m-1}(s-j)^{\sum_{i=j+1}^{2t+r-m}\nu_i}}{(-1)^{\sum_{i=1}^{2t+r-m}\nu_i}R^{\frac{-\sum_{i=1}^{2t+r-m}i\nu_i-m+t}{s}}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}} \\ & \left.\Gamma\left(\frac{2\sum_{i=1}^{2t+r-m}(s-i)\nu_i-2m+2T-2+K(N-1)+2t}{2s}\right)\right\}. \end{aligned} \tag{3.19}$$

Finally the corresponding configuration density results:

Corollary 3.3.5. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1}\otimes I_K, h)$*, with* $\Sigma > 0$*, then the Kotz type III non-isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(\frac{2T+K(N-1)-2}{2s}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right)$$
$$\left\{\sum_{\kappa\in P_{2t+r}}\frac{(2t+r)!\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i}\Gamma\left(\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i+2T-2+K(N-1)+2t}{2s}\right)}{(-1)^{\sum_{i=1}^{2t+r}\nu_i}R^{\frac{-\sum_{i=1}^{2t+r}i\nu_i+t}{s}}\prod_{i=1}^{2t+r}\nu_i!(i!)^{\nu_i}}\right.$$
$$+\sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right]$$
$$\sum_{\kappa\in P_{2t+r-m}}\frac{(2t+r-m)!\prod_{j=0}^{2t+r-m-1}(s-j)^{\sum_{i=j+1}^{2t+r-m}\nu_i}}{(-1)^{\sum_{i=1}^{2t+r-m}\nu_i}R^{\frac{-\sum_{i=1}^{2t+r-m}i\nu_i-m+t}{s}}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$
$$\left.\Gamma\left(\frac{2\sum_{i=1}^{2t+r-m}(s-i)\nu_i-2m+2T-2+K(N-1)+2t}{2s}\right)\right\}. \tag{3.20}$$

Here we need to supply conditions on T and s, as before in remark 3.3.1. First we truncate the series at some t and r, then we propose limits for T and s in such way the Gamma arguments are positive.

In fact the restriction $T \neq 1$ and $s \neq 1$ can be omitted en the above expression and it serve as the full Kotz type configuration density and the cases I and II can be proved trivially, by seen carefully the null derivatives; but we divided the Kotz in the three groups as a didactic way for showing easily the particular cases and their relationships.

Recall that if the above densities exists, then the arguments in gamma's must be positive and this suggest a careful election of the Kotz parameters for doing inference, more over, given the complexity of the expressions it is necessary to truncate the series for some t and r in such way that the parameters T, s and R can be selected.

Remark 3.3.2. Published works involving Kotz shape distributions find the k-th derivative of $h(y) = y^{T-1} \mathrm{e}^{-Ry^s}$, by expanding the exponential as a Maclaurin series, and then deriving the subsequent power; for example Díaz-García and Gutiérrez-Jáimez (2006) write (using their notation and their context)

$$\begin{aligned} h(v) &= \frac{\alpha b^{(2c+kr-2)/2\alpha}\Gamma[(1/2)kr]}{\pi^{kr/2}\Gamma[(2c+kr-2)/2\alpha]} v^{c-1} \exp(-bv^{\alpha}) \\ &= \frac{\alpha b^{(2c+kr-2)/2\alpha}\Gamma[(1/2)kr]}{\pi^{kr/2}\Gamma[(2c+kr-2)/2\alpha]} \sum_{l=0}^{\infty} \frac{(-b)^l v^{\alpha l+c-1}}{l!} \end{aligned}$$

and then

$$h^{2t}(v) = \frac{\alpha b^{(2c+kr-2)/2\alpha}\Gamma[(1/2)kr]}{\pi^{kr/2}\Gamma[(2c+kr-2)/2\alpha]} \sum_{z=0}^{\infty} \frac{(-b)^{z+2t}(\alpha(z+2t)+c-1)_{2t}}{v^{-(\alpha(z+2t)+c-1-2t)(z+2t)!}}. \tag{3.21}$$

First, note that it cannot be used in the general derivation we gave for Kotz configuration density, because the integration and sum can not be interchanged. And second, recall that $(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}$, then the Pochhammer argument in (3.21) must be positive and this restricts importantly their use in the usual papers on Kotz shapes distributions, because they are not valid for all the parameters $r, s > 0$ involved in the matrix variate Kotz definition.

This ratifies the importance of our partitional derivative fashion which leads integration-sum interchange and it is valid for all the Kotz parameters.

3.4 Bessel configuration density

Another elliptical distribution is the so called Bessel distribution, explicitly, the $p \times n$ random matrix is said to have a matrix variate symmetric Bessel distribution with parameters $q, r \in \mathbb{R}$, $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $r > 0$, $q > -\frac{np}{2}$, $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{[\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}]^{\frac{q}{2}} K_q\left(\frac{[\mathrm{tr}(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}]^{\frac{1}{2}}}{r}\right)}{2^{q+np-1}\pi^{\frac{np}{2}} r^{np+q}\Gamma\left(q+\frac{np}{2}\right)|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}}, \tag{3.22}$$

where $K_q(z)$ is the modified Bessel function of the third kind; that is

$$K_q(z) = \frac{\pi}{2}\frac{I_{-q}(z) - I_q(z)}{\sin(q\pi)}, \quad |\arg(z)| < \pi, \quad q \text{ is integer},$$

and

$$I_q(z) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+q+1)}\left(\frac{z}{2}\right)^{q+2k}, \, |z| < \infty, |\arg(z)| < \pi.$$

If $q = 0$ and $r = \frac{\sigma}{\sqrt{2}}$, $\sigma > 0$, this distribution is known as the matrix variate Laplace distribution.

In this case the function h takes the form

$$h(y) = \frac{y^{\frac{q}{2}} K_q\left(\frac{1}{r}y^{\frac{1}{2}}\right)}{2^{q+K(N-1)-1}\pi^{\frac{K(N-1)}{2}} r^{K(N-1)+q}\Gamma\left(q+\frac{K(N-1)}{2}\right)}, \tag{3.23}$$

and the required derivatives for the modified Bessel function are given by

$$K_q^{(k)} = \frac{(-1)^k}{2^k}\sum_{m=0}^{k}\binom{k}{m}K_{q-k+2m}(z).$$

Then, the k-th derivative of (3.23) can be computed by lemma 3.3.1 and after some simplification it results

$$h^{(k)} = \frac{1}{2^{q+K(N-1)-1}\pi^{\frac{K(N-1)}{2}} r^{K(N-1)+q}\Gamma\left(q+\frac{K(N-1)}{2}\right)}$$

$$\sum_{m=0}^{k}\binom{k}{m}\left[\prod_{j=0}^{m-1}\left(\frac{q}{2}-j\right)\right]$$

$$\sum_{\kappa\in P_{k-m}}\frac{(-1)^{k-m}(k-m)!\prod_{j=0}^{k-m-1}\left(\frac{1}{2}-j\right)^{\sum_{i=j+1}^{k-m}\nu_i}}{2^{\sum_{i=1}^{k-m}\nu_i} r^{\sum_{i=1}^{k-m}\nu_i}\prod_{i=1}^{k-m}\nu_i!(i!)^{\nu_i}}$$

$$\sum_{n=0}^{\sum_{i=1}^{k-m}\nu_i}\binom{\sum_{i=1}^{k-m}\nu_i}{n} K_{q-\sum_{i=1}^{k-m}\nu_i+2n}\left(\frac{1}{r}y^{\frac{1}{2}}\right) y^{\sum_{i=1}^{k-m}\left(\frac{1}{2}-i\right)\nu_i+\frac{q}{2}-m}.$$

Thus the integral S in (2.32) can be now computed:

$$S = \int_0^\infty h^{(2t+r)}(y) y^{\frac{K(N-1)}{2}+t-1} dy$$

$$= \frac{1}{\pi^{\frac{K(N-1)}{2}}\Gamma\left(q+\frac{K(N-1)}{2}\right)}$$

$$\sum_{m=0}^{2t+r}\binom{2t+r}{m}\left[\prod_{j=0}^{m-1}\left(\frac{q}{2}-j\right)\right]$$

$$\sum_{\kappa\in P_{2t+r-m}}\frac{(-1)^{2t+r-m}(2t+r-m)!\prod_{j=0}^{2t+r-m-1}\left(\frac{1}{2}-j\right)^{\sum_{i=j+1}^{2t+r-m}\nu_i}}{(2r)^{2\sum_{i=1}^{2t+r-m} i\nu_i+2m-2t}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$

$$\sum_{n=0}^{\sum_{i=1}^{2t+r-m}\nu_i}\binom{\sum_{i=1}^{2t+r-m}\nu_i}{n}$$

$$\Gamma\left(\sum_{i=1}^{2t+r-m}(1-i)\,\nu_i - m + \frac{K(N-1)}{2} + t - n\right)$$

$$\Gamma\left(-\sum_{i=1}^{2t+r-m} i\nu_i + q - m + \frac{K(N-1)}{2} + t + n)\right),$$

The conditions for the existence of S are the same indicated for the Bessel distribution plus the conditions demanded by the derivative and the arguments of the gamma's, which must be positive, see remark 3.3.1, for example.

Thus the configuration density follows from (2.31):

Corollary 3.4.1. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1} \otimes I_K, h)$, *with* $\Sigma > 0$, *then the Bessel non-isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{\frac{Kq}{2}}|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(q+\frac{K(N-1)}{2}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right)$$
$$\sum_{m=0}^{2t+r}\binom{2t+r}{m}\left[\prod_{j=0}^{m-1}\left(\frac{q}{2}-j\right)\right]$$
$$\sum_{\kappa\in P_{2t+r-m}}\frac{(-1)^{2t+r-m}(2t+r-m)!\prod_{j=0}^{2t+r-m-1}\left(\frac{1}{2}-j\right)^{\sum_{i=j+1}^{2t+r-m}\nu_i}}{(2r)^{2\sum_{i=1}^{2t+r-m}i\nu_i+2m-2t}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$
$$\sum_{n=0}^{\sum_{i=1}^{2t+r-m}\nu_i}\binom{\sum_{i=1}^{2t+r-m}\nu_i}{n}$$
$$\Gamma\left(\sum_{i=1}^{2t+r-m}(1-i)\,\nu_i-m+\frac{K(N-1)}{2}+t-n\right)$$
$$\Gamma\left(-\sum_{i=1}^{2t+r-m}i\nu_i+q-m+\frac{K(N-1)}{2}+t+n\right).$$

Finally, if $q = 0$ and $r = \frac{\sigma}{\sqrt{2}}$, $\sigma > 0$ in corollary 3.4.1, then we have the Laplace non-isotropic noncentral configuration density.

Again it is important to note that for doing inference the above series must be truncated and the Bessel parameters chosen in order that the gamma's exist, see remark 3.3.1, for example.

3.5 Jensen-Logistic configuration density

Recall that the $p \times n$ random matrix X is said to have a matrix variate symmetric Jensen-logistic distribution with parameters $M : p \times n$, $\Sigma : p \times p$, $\Phi : n \times n$ with $\Sigma > 0$, and $\Phi > 0$ if its probability density function is

$$\frac{c \operatorname{etr} -(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}}{|\Sigma|^{\frac{n}{2}}|\Phi|^{\frac{p}{2}}\left(1+\operatorname{etr} -(X-M)'\Sigma^{-1}(X-M)\Phi^{-1}\right)^{2}}, \tag{3.24}$$

where

$$c = \frac{\pi^{\frac{np}{2}}}{\Gamma\left(\frac{np}{2}\right)}\int_0^\infty z^{\frac{np}{2}-1}\frac{\mathrm{e}^{-z}}{\left(1+\mathrm{e}^{-z}\right)^2}dz.$$

For this case we put h as

$$h(y) = c\,\mathrm{e}^{-y}\left(1+\mathrm{e}^{-y}\right)^{-2}, \tag{3.25}$$

then the k-th derivative can be performed by using again lemma 3.3.1, and after some simplification we obtain

$$h^{(k)} = c\sum_{m=0}^{k}\binom{k}{m}\sum_{\kappa\in P_{k-m}}\frac{(k-m)!\left(\sum_{i=1}^{k-m}\nu_i+1\right)!\,\mathrm{e}^{-(1+\sum_{i=1}^{k-m}\nu_i)y}}{(-1)^{m+\sum_{i=1}^{k-m}(1+i)\nu_i}\prod_{i=1}^{k-m}\nu_i!(i!)^{\nu_i}\left(1+\mathrm{e}^{-y}\right)^{2+\sum_{i=1}^{k-m}\nu_i}}.$$

So (2.32) becomes

$$\begin{aligned} S &= \int_0^\infty h^{(2t+r)}(y)y^{\frac{K(N-1)}{2}+t-1}dy \\ &= c\sum_{m=0}^{2t+r}\binom{2t+r}{m}\sum_{\kappa\in P_{2t+r-m}}\frac{(2t+r-m)!\left(\sum_{i=1}^{2t+r-m}\nu_i+1\right)!}{(-1)^{m+\sum_{i=1}^{2t+r-m}(1+i)\nu_i}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}} \\ &\int_0^\infty\frac{\mathrm{e}^{-(1+\sum_{i=1}^{2t+r-m}\nu_i)y}}{\left(1+\mathrm{e}^{-y}\right)^{2+\sum_{i=1}^{2t+r-m}\nu_i}}y^{\frac{K(N-1)}{2}+t-1}dy; \end{aligned}$$

and finally, we get the configuration from (2.31):

Corollary 3.5.1. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \Sigma_{N-1\times N-1}\otimes I_K, h)$, with $\Sigma > 0$, then the Jensen-logistic non-isotropic noncentral configuration density is given by*

$$\frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)}\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r$$
$$\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}\right)$$
$$c\sum_{m=0}^{2t+r}\binom{2t+r}{m}\sum_{\kappa\in P_{2t+r-m}}\frac{(2t+r-m)!\left(\sum_{i=1}^{2t+r-m}\nu_i+1\right)!}{(-1)^{m+\sum_{i=1}^{2t+r-m}(1+i)\nu_i}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$
$$\int_0^\infty\frac{\mathrm{e}^{-(1+\sum_{i=1}^{2t+r-m}\nu_i)y}}{(1+\mathrm{e}^{-y})^{2+\sum_{i=1}^{2t+r-m}\nu_i}}y^{\frac{K(N-1)}{2}+t-1}dy \tag{3.26}$$

And by definition of ν_i and t, providing that the integral in c exists, we can see that the above integral also exists, however, for inference and for a meaningful sample of configuration, the above series must be truncated enough and them much of the terms in the derivatives vanish.

Remark 3.5.1. A general configuration density.

We have seen that the classical matrix variate elliptically contoured distributions proposed in the generalized multivariate analysis (Pearson, Kotz, Bessel, Jensen-Logistic) can be studied with our simplification formula for the configuration density, by expressing the function h as $h(y) = s(y)g(f(y))$, so, lemma 3.3.1 can be applied for the required derivative in (2.32). However we can take a more general form of h, that is $h(y) = s(r(y))g(f(y))$ which also can be derived by lemma 3.3.1 and, if the normalization constant of the distribution has a closed form, then the integral can be computed.

Clearly the integral S in (2.32) cannot be computed in general, but provided its finiteness, then the corresponding elliptically configuration can exists under certain parameter conditions.

Chapter 4

Inference and Applications

In the preceding chapters, we derived the noncentral configuration density under an elliptical model and by using partition theory, a number of explicit configuration densities were obtained; i.e. configuration densities associated with the matrix variate symmetric Kotz type distributions (it includes normal), the matrix variate Pearson type VII distributions (it includes t and Cauchy distributions), the matrix variate symmetric Bessel distribution (it includes Laplace distribution) and the matrix variate symmetric Jensen-logistic distribution. The configuration density of any elliptical model was set in terms of zonal polynomials which now can be efficiently computed by Koev and Edelman (2006), and in consequence, the inference problem can be studied and solved with the exact densities instead of usual constraints and asymptotic distributions, and approximations of the statistical shape theory works (see Goodall and Mardia (1993), Dryden and Mardia (1998) and the references there in). The general procedure becomes very clear now and the underlying problem, the programming problem, is simply time consuming.

Thus two perspectives can be explored, first, the inference based on exact distributions and second, their applications in shape theory.

The general procedure for performing inference of any elliptical model is proposed and it is set in such manner that the published efficient numerical algorithms for confluent infinite series type involving zonal polynomials, can be used; this is outlined in section 4.1.

More over, a further simplification of the closed computational problem is also proposed, the study of finite configuration densities (section 4.2); a subfamily of them is derived and as a simple example of their use, exact inference for testing configuration location differences in some applied problems are provided in section 4.3. The applications involve Biology (mouse vertebra, gorilla skulls, girl and boy craniofacial studies), Medicine (brain MR scans of schizophrenic patients) and image analysis (postcode recognition).

4.1 Inference for elliptical configuration models

Our proposal is to use the elliptically contoured distribution to model population configurations (2.31) for some particular cases. For this, we consider a random sample of n independent and identically distributed observations $U_1, \ldots, U_n$ obtained from

$$Y_i \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h), \quad i = 1, \ldots, n,$$

by mean of (2.27).

Now we define the configuration population parameters. Let $CD(U; \mathcal{U}, \sigma^2)$ be the exact configuration density, where $\mathcal{U}$ is the location parameter matrix of the configuration population (we just say configuration location) and σ^2 is population scale parameter. Both $\mathcal{U}$ and σ^2 are the parameters to estimate. More exactly, let $\mu \neq 0_{N-1\times K}$ be the parameter matrix of the elliptical density Y considered in theorem 2.2.1; if we write it as $\mu = (\mu_1' \mid \mu_2')'$, where $\mu_1 : K \times K$ (nonsingular) and $\mu_2 : q = N - K - 1 \geq 1 \times K$, then, according to (2.27), we can define the configuration location parameter matrix $\mathcal{U} : N-1 \times K$ as follows: $\mathcal{U} = (I_K \mid \mathcal{V}')'$ where $\mathcal{V} = \mu_2\mu_1^{-1}$; and $\mathcal{V} : q = N - K - 1 \geq 1 \times K$ contains $q \times K$ configuration location parameters to estimate. Then, taking into account this remark and using the same notation of Dryden and Mardia (1998), p. 144-145. we have:

$$\log L(U_1, \ldots, U_n; \mathcal{V}, \sigma^2) = \sum_{i=1}^{n} \log CD(U_i; \mathcal{V}, \sigma^2).$$

Finally, the maximum likelihood estimators for location and scale parameters are

$$(\widetilde{\mathcal{V}}, \widetilde{\sigma^2}) = \arg \sup_{\mathcal{V}, \sigma^2} \log L(U_1, \ldots, U_n; \mathcal{V}, \sigma^2). \tag{4.1}$$

The general configuration density is an infinite series in the eigenvalues of the zonal polynomial argument, and as usual in other shape densities involving those polynomials (see Goodall and Mardia (1993)), the exact likelihood estimators cannot be found, at present, in an exact form, then we need to perform a numerical optimization. Fortunately, the classical elliptical configuration densities can be efficiently computable (see step III below) and the likelihood maximization procedure is computationally possible. Now, for the numerical optimization we can use a number of routines, which, clearly, are based on the initial point for estimation. In our case, consider the Helmertized landmark data $Y_i \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$ $i = 1, \ldots, n$ (see (2.27)) and let

$\widetilde{\mu} = (\widetilde{\mu}_1' \mid \widetilde{\mu}_2')'$ and $\tilde{\sigma}^2$ be the maximum likelihood estimators of the location parameter matrix $\mu_{N-1\times K}$ and the scale parameter σ^2 of the elliptical distribution under consideration, so, given that

$$U_i'\Sigma^{-1}\mu\mu'\Sigma^{-1}U_i(U_i'\Sigma^{-1}U_i)^{-1} = Y_i'\Sigma^{-1}\mu\mu'\Sigma^{-1}Y_i(Y_i'\Sigma^{-1}Y_i)^{-1},$$

then an initial point can be $x_0 = (\mathrm{vec}'(\mathcal{V}_0'), \sigma_0^2)$, where $\mathcal{V}_0 = \widetilde{\mu}_2\widetilde{\mu}_1^{-1}$ and $\sigma_0^2 = \tilde{\sigma}^2$.

Now, we propose the directions for solving the inference in the next few steps.

4.1.1 Step I. Families of isotropic elliptical configuration densities

A first step considers a list of configuration densities. We just write down below the whole group of classical densities in the matrix case studied by standard books in matrix elliptically contoured distributions such as Gupta and Varga (1993). But we must note that any function $h(\cdot)$ which satisfies (2.32) and the conditions of theorem 2.2.1 can be appropriate.

Most of the applications in statistical theory of shape reside on the isotropic model (see Dryden and Mardia (1998)), so in the case of the noncentral elliptical configuration density if we take $\Sigma = \sigma^2 I_{N-1}$ in the preceding non-isotropic densities we get a list of suitable distributions for inference; which, we noted are expanded in terms of zonal polynomials and they can be computed, after an efficient method of Edelman's group (see Koev and Edelman (2006) and Dumitriu *et al* (2004)). Note, that we can consider a more enriched structure, for example $\Sigma = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N-1}^2)$ (which suppose a different variance in each landmark component), and similar diagonal structures.

In the isotropic case we have the following list of the densities here derived

Person type VII configuration density

Corollary 4.1.1. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$*, then the isotropic noncentral Pearson type VII configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$ {}_1P_1\left(\left(s - \frac{K(N-1)}{2}\right)_t \left(1 + \frac{\operatorname{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)}{R}\right)^{-s+\frac{K(N-1)}{2}-t} : \right.$$
$$\left. \frac{N-1}{2}; \frac{K}{2}; \frac{1}{R\sigma^2}U'\mu\mu'U(U'U)^{-1}\right), \tag{4.2}$$

where

$$ {}_1P_1(f(t) : a; b; X) = \sum_{t=0}^{\infty} \frac{f(t)}{t!} \sum_{\tau} \frac{(a)_\tau}{(b)_\tau} C_\tau(X). \tag{4.3}$$

Corollary 4.1.2. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$*, then the isotropic noncentral* t *configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$ {}_1P_1\left(\left(\frac{R}{2}\right)_t \left(1 + \frac{\operatorname{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)}{R}\right)^{-\frac{R}{2}-t} : \right.$$
$$\left. \frac{N-1}{2}; \frac{K}{2}; \frac{1}{R\sigma^2}U'\mu\mu'U(U'U)^{-1}\right). \tag{4.4}$$

Corollary 4.1.3. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$*, then the isotropic noncentral Cauchy configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$$ {}_1P_1\left(\left(\frac{1}{2}\right)_t \left(1 + \operatorname{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)\right)^{-\frac{1}{2}-t} : \right.$$
$$\left. \frac{N-1}{2}; \frac{K}{2}; \frac{1}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right). \tag{4.5}$$

Kotz type configuration density

Recall that the series must be truncated at some t and r, then the limits for the parameters T and s are selected in such way that the Gamma's arguments are positive, see remark 3.3.1.

Corollary 4.1.4. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$, then the Kotz type I isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)} \sum_{t=0}^{\infty} \frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(T-1+\frac{K(N-1)}{2}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(-\frac{R}{\sigma^2}\mu'\mu\right)\right]^r \sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau} C_\tau\left(\frac{R}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right)$$
$$\left\{\Gamma\left(T-1+\frac{K(N-1)}{2}+t\right)\right. \tag{4.6}$$
$$\left. + \sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right](-1)^m\Gamma\left(T-1-m+\frac{K(N-1)}{2}+t\right)\right\}.$$

Corollary 4.1.5. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$ and $T = 1$, then the Kotz type I isotropic noncentral configuration density simplifies to*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)\operatorname{etr}\left(\frac{R}{\sigma^2}\mu'U(U'U)^{-1}U'\mu - \frac{R}{\sigma^2}\mu'\mu\right)}{\pi^{Kq/2}|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$${}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{R}{\sigma^2}\mu'U(U'U)^{-1}U'\mu\right). \tag{4.7}$$

Corollary 4.1.6. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1} \otimes I_K, h)$, $T = 1$, $R = \frac{1}{2}$ then the Kotz type I isotropic noncentral configuration density simplifies to normal configuration density and it is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)\operatorname{etr}\left(\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu - \frac{1}{2\sigma^2}\mu'\mu\right)}{\pi^{Kq/2}|I_K + V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}$$
$${}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu\right). \tag{4.8}$$

Corollary 4.1.7. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1}\otimes I_K, h)$, then the Kotz type II isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K+V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(\frac{K(N-1)}{2s}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\mathrm{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right)$$
$$\sum_{\kappa\in P_{2t+r}}\frac{(2t+r)!\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i}\Gamma\left(\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i+K(N-1)+2t}{2s}\right)}{(-1)^{\sum_{i=1}^{2t+r}\nu_i}R^{-\frac{\sum_{i=1}^{2t+r}i\nu_i-t}{s}}\prod_{i=1}^{2t+r}\nu_i!(i!)^{\nu_i}}. \tag{4.9}$$

Corollary 4.1.8. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1}\otimes I_K, h)$, then the Kotz type III isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K+V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{\Gamma\left(\frac{K(N-1)}{2}\right)}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(\frac{2T+K(N-1)-2}{2s}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\mathrm{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right)$$
$$\left\{\sum_{\kappa\in P_{2t+r}}\frac{(2t+r)!\prod_{j=0}^{2t+r-1}(s-j)^{\sum_{i=j+1}^{2t+r}\nu_i}\Gamma\left(\frac{2\sum_{i=1}^{2t+r}(s-i)\nu_i+2T-2+K(N-1)+2t}{2s}\right)}{(-1)^{\sum_{i=1}^{2t+r}\nu_i}R^{\frac{-\sum_{i=1}^{2t+r}i\nu_i+t}{s}}\prod_{i=1}^{2t+r}\nu_i!(i!)^{\nu_i}}\right.$$
$$+\sum_{m=1}^{2t+r}\binom{2t+r}{m}\left[\prod_{i=0}^{m-1}(T-1-i)\right]$$
$$\sum_{\kappa\in P_{2t+r-m}}\frac{(2t+r-m)!\prod_{j=0}^{2t+r-m-1}(s-j)^{\sum_{i=j+1}^{2t+r-m}\nu_i}}{(-1)^{\sum_{i=1}^{2t+r-m}\nu_i}R^{\frac{-\sum_{i=1}^{2t+r-m}i\nu_i-m+t}{s}}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$
$$\left.\Gamma\left(\frac{2\sum_{i=1}^{2t+r-m}(s-i)\nu_i-2m+2T-2+K(N-1)+2t}{2s}\right)\right\}. \tag{4.10}$$

Bessel configuration density

Corollary 4.1.9. *If* $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1}\otimes I_K, h)$*, then the Bessel isotropic noncentral configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{\frac{Kq}{2}}|I_K+V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)\Gamma\left(q+\frac{K(N-1)}{2}\right)}$$
$$\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)\right]^r\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right)$$
$$\sum_{m=0}^{2t+r}\binom{2t+r}{m}\left[\prod_{j=0}^{m-1}\left(\frac{q}{2}-j\right)\right]$$
$$\sum_{\kappa\in P_{2t+r-m}}\frac{(-1)^{2t+r-m}(2t+r-m)!\prod_{j=0}^{2t+r-m-1}\left(\frac{1}{2}-j\right)^{\sum_{i=j+1}^{2t+r-m}\nu_i}}{(2r)^{2\sum_{i=1}^{2t+r-m}i\nu_i+2m-2t}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$
$$\sum_{n=0}^{\sum_{i=1}^{2t+r-m}\nu_i}\binom{\sum_{i=1}^{2t+r-m}\nu_i}{n}$$
$$\Gamma\left(\sum_{i=1}^{2t+r-m}(1-i)\,\nu_i-m+\frac{K(N-1)}{2}+t-n\right)$$
$$\Gamma\left(-\sum_{i=1}^{2t+r-m}i\nu_i+q-m+\frac{K(N-1)}{2}+t+n\right).$$

Recall that the positive Gamma arguments are supplied after the truncation of the series at some t and r, then the bounds for the remaining parameters can be obtained.

Finally, if $q=0$ and $r=\frac{\sigma}{\sqrt{2}}$, $\sigma>0$ in corollary 4.1.9, then we have the Laplace isotropic noncentral configuration density.

Jensen-Logistic configuration density

Corollary 4.1.10. *If $Y \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1}\otimes I_K, h)$, then the Jensen-logistic isotropic noncentral configuration density is given by*

$$\frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|I_K+V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\sum_{t=0}^{\infty}\frac{1}{t!\Gamma\left(\frac{K(N-1)}{2}+t\right)}\sum_{r=0}^{\infty}\frac{1}{r!}\left[\operatorname{tr}\left(\frac{1}{\sigma^2}\mu'\mu\right)\right]^r$$
$$\sum_{\tau}\frac{\left(\frac{N-1}{2}\right)_\tau}{\left(\frac{K}{2}\right)_\tau}C_\tau\left(\frac{1}{\sigma^2}U'\mu\mu'U(U'U)^{-1}\right)$$
$$c\sum_{m=0}^{2t+r}\binom{2t+r}{m}\sum_{\kappa\in P_{2t+r-m}}\frac{(2t+r-m)!\left(\sum_{i=1}^{2t+r-m}\nu_i+1\right)!}{(-1)^{m+\sum_{i=1}^{2t+r-m}(1+i)\nu_i}\prod_{i=1}^{2t+r-m}\nu_i!(i!)^{\nu_i}}$$
$$\int_0^\infty\frac{\mathrm{e}^{-(1+\sum_{i=1}^{2t+r-m}\nu_i)y}}{(1+\mathrm{e}^{-y})^{2+\sum_{i=1}^{2t+r-m}\nu_i}}y^{\frac{K(N-1)}{2}+t-1}dy. \tag{4.11}$$

4.1.2 Step II. Choosing the elliptical configuration density

Here we have the main advantage of working with elliptical models, the possibility of choosing a distribution for the landmark data, recall that the normal assumption is constantly repeated in the statistical shape theory context. Recall that the main assumptions for inference in this works are supported by independent and identically elliptically contoured distributed observations

$$Y_i \sim E_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1}\otimes I_K, h),\ i=1,\dots,n.$$

According to our assumptions we can consider Schwarz (1978) as an appropriate technique for choosing the elliptical model. Explicitly, the procedure is as follows: consider k elliptical models, then perform the maximization of the likelihood function separately for each model $j = 1,\dots,k$, obtaining say,

$M_j(Y_1, \ldots, Y_n)$, then Schwarz's criterion for a large-sample is given by

$$\textit{Choose the model for which } \log M_j(Y_1, \ldots, Y_n) - \frac{1}{2} k_j \log n \textit{ is largest},$$

where k_j is the dimension (number of parameters) of the model j.

Remark 4.1.1. The preceding result can be implemented for choosing a shape model, i.e. given an independent and identically distributed random sample of landmark data and a list of shape distributions: pre-shape, size and shape, shape, reflection shape, reflection size and shape, cone, disk, (all of them supported by Euclidean transformations), configuration (supported by affine transformations), and projection, etc. we can select the best shape-transformation-model. However, it is constrained by the computation of the densities, and as we can check statistical shape literature, the Euclidean based shape densities have important difficulties for computations even in the gaussian case (most of them have no an elliptical version yet), see Goodall and Mardia (1993), Dryden and Mardia (1998) and the references there in, but it is not the case with our configuration densities. We will let these comparisons for a subsequent work.

4.1.3 Step III. Configuration Location

Once the elliptical model is selected we find the estimators of configuration location and scale parameters by mean of (4.1). The crucial point here is the computation of the configuration density; If the selected model is the Gaussian one, then the matlab algorithms for confluent hypergeometric functions of matrix argument by Koev and Edelman (2006) gives the solution very efficiently, this solves in fact the inference problem proposed by Goodall and Mardia (1993), corrected by Díaz-García *et al.* (2003) and corrected again here. We highlight that the cited computation of the ${}_1F_1(a;c;X)$ series restricted to the truncation and it is an open problem addressed in the last section of Koev and Edelman (2006), however the fast algorithms let a sort of numerical experiments until a given precision is reached, so the optimization problem remains in terms of the truncation and the set precision, but this occurs, clearly, since it is an intrinsic problem of any numerical optimization problem.

But, if the selected model is not Gaussian, we could think that the problems remains open, but fortunately, the configuration densities can be computed efficiently by using the same work of Koev and Edelman (2006).

First, represent the configuration density in theorem 2.2.1 by using the notation ${}_1P_1(f(t):a;c;X)$ introduced in corollary 3.2.1, (see (3.5)) i.e.

$$ {}_1P_1(f(t):a;c;X) = \sum_{t=0}^{\infty} \frac{f(t)}{t!} \sum_{\tau} \frac{(a)_\tau}{(c)_\tau} C_\tau(X), $$

thus the configuration density is written as

$$ A \; {}_1P_1(f(t):a;c;X), \tag{4.12} $$

where

$$ A = \frac{\pi^{K^2/2}\Gamma_K\left(\frac{N-1}{2}\right)}{|\Sigma|^{\frac{K}{2}}|U'\Sigma^{-1}U|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}, \tag{4.13} $$

$$ f(t) = \sum_{r=0}^{\infty} \frac{\left[\operatorname{tr}\left(\mu'\Sigma^{-1}\mu\right)\right]^r}{r!\Gamma\left(\frac{K(N-1)}{2}+t\right)} \int_0^{\infty} h^{(2t+r)}(y) y^{\frac{K(N-1)}{2}+t-1} dy, \tag{4.14} $$

$$ a = \frac{N-1}{2}, \quad c = \frac{K}{2}, \tag{4.15} $$

and

$$ X = U'\Sigma^{-1}\mu\mu'\Sigma^{-1}U(U'\Sigma^{-1}U)^{-1}. \tag{4.16} $$

Note that f depends on r, but we are interested in the whole series indexed by t, however the r dependency should be clear in the context. And given that the integral S is finite then the series $A\ {}_1P_1(f(t) : a; b; X)$, the configuration, is a density and thus ${}_1P_1(\cdot)$ converges, and the series $f(t)$ also converges in the respective domain given a the particular elliptical distribution is considered.

Unfortunately, the configuration density $A\ {}_1P_1(f(t) : a; c; X)$ is an infinite series, given that $a = \frac{N-1}{2}$ and $c = \frac{K}{2}$ are positive. (recall that N is the number of landmarks, K is de dimension and $N - K - 1 \geq 1$). So a truncation is needed if we want to use it directly by using computation of zonal polynomials.

Expression (4.12) belongs to the class of eq. (1.6), p.3 from Koev and Edelman (2006), and as they affirm "With minimal changes our algorithms (for hypergeometric functions of one matrix arguments) can approximate the hypergeometric function of two matrix arguments..., and more generally functions of the form $G(x) = \sum_{k=0}^{\infty} \sum_{\kappa} a_\kappa C_\kappa^\alpha(X)$, for arbitrary coefficients a_κ at a similar computational cost," see, eq. (6.5) of Koev and Edelman (2006), and they add "Although the expression (6.5) is not a hypergeometric function of a matrix argument, its truncation for $|\kappa| \leq m$ has the form (1.6), and is computed analogously."

Then, in principle, the configuration densities can be evaluated efficiently with the fast algorithms of Koev and Edelman (2006) and the corresponding inferences problem can be solved numerically. And at this stage, by using for example the compatible matlab routine $fminsearch$ with the modified matlab files of Koev and Edelman (2006), we have the estimators for the configuration location and the scale parameters of the "best" elliptical model chosen with Schwarz's criterion. We arrive then, to the final step.

4.1.4 Step IV. Hypothesis testing

Finally, given that the likelihood can be evaluated and optimized, then a sort of likelihood ratio tests can be performed for testing a particular configuration for a population, or testing for differences in configuration between two populations, or testing one-dimensional uniform shear of two populations, etc.

In the statistical shape analysis, the large sample standard likelihood ratio tests are the most frequently used, see for example Dryden and Mardia (1998), by mean of Wilk's theorem.

Explicitly, for testing whether

$$H_0 : \mathcal{U} \in \Omega_0$$

versus

$$H_a : \mathcal{U} \in \Omega_1,$$

where $\Omega_0 \subset \Omega_1 \subseteq \mathbb{R}^{Kq}$, with $\dim(\Omega_0) = p < Kq$ and $\dim(\Omega_a) = r \leq Kq$. Thus, the -2 log-likelihood ratio is given by

$$-2 \log \Lambda = 2 \sup_{H_a} \log L(\mathcal{U}, \sigma^2) - 2 \sup_{H_0} \log L(\mathcal{U}, \sigma^2)$$

and by the Wilk's theorem for large samples, the distribution of the null hypothesis H_0 obeys

$$-2 \log \Lambda \approx \chi^2_{r-p},$$

see Dryden and Mardia (1998).

In a similar way we can test differences in configuration between two populations, etc.

Suppose that the last hypothesis is rejected, then an interesting test can be performed one-dimensional uniform shear of two populations which determines the amount of deformation axes by axes. Note that the classical statistical shape analysis (pre-shape, size and shape, shape, reflection shape, reflection size and shape, cone, disk,) which is based on Euclidean transformations assume that any shape is uniform deformed in any dimension, which certainly is very idealistic, but the configuration density accept different uniform shearing among the axes.

Explicitly, if we want to test uniform shear in the i coordinate of two populations, then the testing procedure lies on $H_0 : \mu_1 B = \mu_1 B$ versus $H_0 : \mu_1 B \neq \mu_1 B$, where $B = (0, \ldots, i, \ldots, 0)'$ and the configuration density U goes to UB. Note that the new configuration density is simpler, since it is just a vector density and it is easier of computing.

Thus, the whole inference procedure of the above four steps can be carried out for a particular landmark data, and up here we can consider our thesis problem solved.

4.2 Further simplifications: finite configuration densities

Even though the whole elliptical configuration problem is clear, there are interesting simplifications which open promissory future work. We explore the problem in this section, before ending this work with some applications.

Consider now the complex problem of the estimation of the configuration location; as we highlighted before, no published works using and exact non central density expressed in terms of zonal polynomials is available, only the asymptotic expansions and non feasible assertions on variance have been use for testing mean shape differences in other contexts, the ratification of this idea revealed in every application of the excellent and summary work on shape theory by Dryden and Mardia (1998). So any improvement of this problem by working with the exact density can be considered a good achievement in a theory which sums more than 50 years of probability density functions expanded by zonal polynomials.

In the context of shape theory for example, the general size-and-shape, and the isotropic gaussian configuration densities (Goodall and Mardia (1993) and Díaz-García *et al.* (2003)) are expanded in terms of zonal polynomials, but at the time of the first appearances of these distributions, no accurate numerical methods for large degrees of the polynomials were known. This forced to use approximations of the hypergeometric functions to do the inference, but even by using these approaches, the inference problem has been studied in the isotropic case and only with $\sigma \to 0$, see Goodall and Mardia (1993). Only, very recently, Koev and Edelman (2006) solved the problem of computation of zonal and hypergeometric type functions, as we described in the last section, and clearly the efficient computation is the key for doing inference, but only 2 years ago the solution of numerical computation for those polynomials was given by Edelman's gruop (the general formulae for any degree remains open, only up second order is solved for general Jack polynomials, which include zonal polynomials, see Caro-Lopera *et al* (2007)).

However, as the zonal polynomials are computable very fast, the problem now resides in the convergence and the truncation of the above series for performing the numerical optimization. In fact, in the same reference of Koev and Edelman (2006) we read:

"Several problems remain open, among them automatic detection of convergence .... and it is unclear how to tell when convergence sets in. Another open problem is to determine the best way to truncate the series. "

Thus the implicit numerical difficulties for truncation of any configuration density series of type (4.12) motivates two areas of investigation: one, continue the numerical approach started by (Koev and Edelman (2006)) with the confluent hypergeometric functions and extend it to the case of some configuration series type Kotz, Pearson, Bessel, Jensen-Logistic, for example; or second, propose a theoretical approach for solving the problem analytically.

In the next few lines we establish the second question and leave their implications for future work.

First represent the configuration density in theorem 2.2.1 by using the notation ${}_1P_1(f(t):a;c;X)$, as it was done in (4.12)-(4.16).

Unfortunately, the configuration density $A\ {}_1P_1(f(t):a;c;X)$ is an infinite series, given that $a=\frac{N-1}{2}$ and $c=\frac{K}{2}$ are positive. (recall that N is the number of landmarks, K is de dimension and $N-K-1\geq 1$). So a truncation is needed as we explained in the step III of the general inference procedure.

But the above series can be finite if we use the following basic principle.

Lemma 4.2.1. *Let $N-K-1 \geq 1$ as usual, and consider the infinite configuration density*

$$CD_1 = A \ {}_1P_1 \left(f(t) : \frac{N-1}{2}; \frac{K}{2}; X \right).$$

If the dimension K is even (odd) and the number of landmarks N is odd (even), respectively, then the equivalent configuration density

$$CD_2 = A \ {}_1P_1 \left(g(t) : -\left(\frac{N-1}{2} - \frac{K}{2} \right); \frac{K}{2}; h(X) \right)$$

is a polynomial of degree $K \left(\frac{N-1}{2} - \frac{K}{2} \right)$ in the latent roots of the matrix X (otherwise the series is infinite).

By equivalent configuration we just mean that $CD_1 = CD_2$, and the functions $g(t)$ and $h(X)$ depends on the function $f(t)$ under consideration, and which can be established by using the integral representation of CD_1.

Proof. Recall that $\tau = (t_1, \ldots, t_K)$, $t_1 \geq t_2 \geq \cdots t_K \geq 0$, is a partition of t and

$$(\alpha)_\tau = \prod_{i=1}^{K} \left(\alpha - \frac{1}{2}(i-1) \right)_{t_i},$$

where

$$(\alpha)_t = \alpha(\alpha+1)\cdots(\alpha+t-1), \quad (\alpha)_0 = 1.$$

Now, If K is even (odd) and N is odd (even) then $-\left(\frac{N-1}{2} - \frac{K}{2}\right) = -\frac{q}{2}$ is a negative integer and clearly $(-\frac{q}{2})_\tau = 0$ for every $t \geq \frac{Kq}{2} + 1$, then CD_2 is a polynomial of degree $\frac{Kq}{2}$ in the latent roots of X. ■

So, the addressed truncation problem of an infinite configuration density can be solved by finding an equivalent finite configuration density according to the preceding lemma and selecting an appropriate number of landmarks in the figure.

Given an elliptical configuration density CD_1 indexed by function $f(t)$, $a = \frac{N-1}{2} > 0$,$c = \frac{K}{2} > 0$, the crucial point consists of finding an integral representation valid for $c-a = -\frac{q}{2} < 0$ leading an equivalent elliptical configuration density CD_2 indexed by some function $g(t)$. Then the finiteness of CD_2 follows from K even (odd) and N odd (even), respectively.

We already saw a type of these relations, when $f(t)$ is a constant, i.e. in corollary 4.1.5; in this case the main principle is reduced to the Kummer relations; and the corresponding configuration densities (which include Gaussian) are finite by selecting an odd (even) number of landmarks N according to an even (odd) dimension K, respectively. The implications of the finiteness for applications will avoid the addressed open problem for truncation proposed in Koev and Edelman (2006).

The relations (3.2) were studied by Herz in the 50's and constitute the so called Kummer relations as a generalization of the known scalar case. In this sense, we can see the relations in lemma 4.2.1 as the the generalized Kummer relations associated to the elliptical based non constant function $f(t)$.

Perhaps a comment about the proof of (3.2) and their use in the normal configuration density is convenient. Herz (1955) (eq. (2.8), p.488) proved that

$$ {}_1F_1(a;c;X) = \operatorname{etr} X {}_1F_1(c-a;c;-X) $$

is valid for all $\operatorname{Re} c > (K+1)/2 - 1$ by using a Laplace integral representation of ${}_1F_1(a;c;X)$, and clearly it admits $c - a < 0$ which is the key point for the finiteness of the normal configuration density here derived (see also Goodall and Mardia (1993), Díaz-García *et al.* (2003)). However there is another representation of confluent hypergeometric, a beta type integral, which leads the Kummer relations, also proved by Herz (1955) (eq. (2.9), p.488), but it is not true for our configuration density requirements $c-a<0$; in particular, Muirhead (1982), p. 265. Th. 7.4.3, proves Kummer relations by using the last representation (Muirhead (1982), p. 264. Th. 7.4.2), but precisely Díaz-García *et al.* (2003), p. 143, cite Muirhead's th. 7.4.3 for the gaussian configuration density simplification, which is not valid for $c-a<0$, the right citation should be Herz (1955) (eq. (2.8), p.488, which uses the Laplace representation.

The above discussion it is important for generalization of Kummer relations; for example, the author has not knowledge of this relations for non constant $f(t)$, i.e. expressions of $g(t)$ and $h(X)$ for non normal models. Some advances in this direction are available from the author, for example, the generalized Kummer relations for a Kotz type I (T positive integer), and a Pearson type VII, based on a Beta type integral representation; which have ratified that ${}_1P_1\left(f(t):a;c;X\right) ={}_1 P_1\left(g(t):c-a;c;-X\right)$, for the corresponding f,g, but in the case of $c-a>0$. The next step is to prove the relations for $c-a<0$, by a Laplace representation type, then lemma 4.2.1 can be applied to Kotz type and Pearson type VII configuration densities and the respective series become finite. However, the addressed last problem requires the evaluation of certain new integrals involving invariant polynomials (Davis (1980)), so the first stage relations are good motivations for continuing to study the referred generalized relations.

Meanwhile, fortunately, we can performed inference with finite series (3.15) specially with the Gaussian case $R=\frac{1}{2}$.

Corollary 4.2.1. *If* $Y \sim \mathcal{N}_{N-1\times K}(\mu_{N-1\times K}, \sigma^2 I_{N-1}\otimes I_K)$*,* K *is even (odd) and* N *is odd (even), respectively, then the finite isotropic noncentral normal* *configuration density is given by*

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2}|I_K+V'V|^{\frac{N-1}{2}}\Gamma_K\left(\frac{K}{2}\right)}\operatorname{etr}\left(\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu-\frac{1}{2\sigma^2}\mu'\mu\right)$$
$$ {}_1F_1\left(-\frac{q}{2};\frac{K}{2};-\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu\right), \tag{4.17}$$

and it is a polynomial of degree $K\left(\frac{N-1}{2}-\frac{K}{2}\right)$
in the latent roots of $\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu$.

4.3 Applications

In this section, we consider planar classical applications in statistical shape analysis. The following situations are sufficiently studied by shape based on Euclidian transformations and asymptotic formulae. We will use here exact inference in the sense that we will use the exact densities and compute the likelihood exactly by using zonal polynomial theory.

We will test configuration differences under the exact gaussian configuration density, and the applications include Biology (mouse vertebra, gorilla skulls, girl and boy craniofacial studies), Medicine (brain MR scans of schizophrenic patients) and image analysis (postcode recognition).

According to the preceding section, we perform exact inference based on the finite normal configuration density and a sort of landmark data usually studied in the context of normal shape distributions.

First we start with the two dimensional case, then corollary 4.2.1 turns:

Corollary 4.3.1. *If* $Y\sim\mathcal{N}_{N-1\times 2}(\mu_{N-1\times 2},\sigma^2 I_{N-1}\otimes I_2)$*, and* N *is odd , then the finite two dimensional isotropic noncentral normal configuration density is* *given by*

$$\frac{\Gamma_2\left(\frac{N-1}{2}\right)}{\pi^{N-3}|I_2+V'V|^{\frac{N-1}{2}}\Gamma_K(1)}\operatorname{etr}\left(\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu-\frac{1}{2\sigma^2}\mu'\mu\right)$$
$$ {}_1F_1\left(-\frac{N-3}{2};1;-\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu\right), \tag{4.18}$$

and it is a polynomial of degree $N-3$ in the two latent roots of $\frac{1}{2\sigma^2}\mu'U(U'U)^{-1}U'\mu$.

Given that most of the applications in shape theory comes from two dimensional images (see Dryden and Mardia (1998)), then it is important to give explicit expressions for the finite series when $N=5,7,9,\ldots$ is small. Let x,y be the eigenvalues of Ω, then we have for $N=5,7,\ldots,21$ the following polynomials of degree $Kq/2=N-3$ in the eigenvalues x,y of Ω; expressions useful for exact inference of the corresponding configuration densities. We use in this case exact formulae for zonal polynomials given by James (1968) see also Caro-Lopera *et al* (2007). In fact all the applications studied in Dryden and Mardia (1998) have maximum 21 landmarks (which supposes a polynomial of 18 degree in the two eigenvalues of corresponding matrix), so the following confluent hypergeometric expressions are sufficient for their corresponding configuration analysis. Note that the cited applications demand formulae for zonal polynomials of second order up maximum twenty degree, and this expressions are available since 60's, so the numerical algorithms of Koev and Edelman (2006) very useful for infinite series but with the addressed problem of truncations, are not needed here and the exact inference on configuration densities historically could be studied since they were proposed by Goodall and Mardia (1993).

Note that the selection of an odd number of landmarks for planar applications suggest deleting one of them of the available tables usually studied for approximations methods, clearly it is also possible to reduce in one, any group of preset even landmark, however we leave the decision to an expert. According to the number of odd landmarks, we suggest some problems studied by Dryden and Mardia (1998) but in the context of finite gaussian configuration densities (we put in parenthesis the original number of landmarks studied by Dryden and Mardia (1998)).

The involved series up 15 landmarks are easily computed as (see section 1.19):

- $N = 5$: Mouse vertebra (6),

$$1 + y + x + 2\,yx \tag{4.19}$$

- $N = 7$: Gorilla skulls (8),

$$1 + 2\,y + 2\,x + \frac{1}{2}\,y^2 + 7\,yx + \frac{1}{2}\,x^2 + 2\,y^2x + 2\,yx^2 + \frac{2}{3}\,y^2x^2 \tag{4.20}$$

- $N = 9$:

$$\begin{aligned}&1 + 3\,y + 3\,x + \frac{3}{2}\,y^2 + 15\,yx + \frac{3}{2}\,x^2 + \frac{1}{6}\,y^3 + \frac{17}{2}\,y^2x\\&+\frac{17}{2}\,yx^2 + \frac{1}{6}\,x^3 + y^3x + \frac{16}{3}\,y^2x^2 + yx^3 + \frac{2}{3}\,y^3x^2\\&+\frac{2}{3}\,y^2x^3 + \frac{4}{45}\,y^3x^3\end{aligned} \tag{4.21}$$

- $N = 11$: Sooty mangabeys (12).

$$\begin{aligned}&1 + 4\,x + 22\,y^2x + 4\,y + \frac{81}{4}\,y^2x^2 + 5\,y^2x^3 + 5\,y^3x^2 + 22\,yx^2\\&+\frac{31}{6}\,y^3x + \frac{31}{6}\,yx^3 + 26\,yx + \frac{1}{24}\,x^4 + \frac{1}{3}\,yx^4 + \frac{1}{3}\,y^4x\\&+\frac{2}{315}\,y^4x^4 + \frac{4}{45}\,y^4x^3 + \frac{4}{45}\,y^3x^4 + \frac{1}{3}\,y^4x^2 + \frac{58}{45}\,y^3x^3\\&+\frac{1}{3}\,y^2x^4 + 3\,y^2 + 3\,x^2 + \frac{2}{3}\,y^3 + \frac{2}{3}\,x^3 + \frac{1}{24}\,y^4\end{aligned} \tag{4.22}$$

- $N = 13$: Brain MR scans of schizophrenic patients (13), postcode recognition (13)

$$
\begin{aligned}
&1 + 5\,x + 45\,y^2x + 5\,y + \frac{655}{12}\,y^2x^2 + \frac{241}{12}\,y^2x^3 + \frac{241}{12}\,y^3x^2 + 45\,yx^2\\
&+\frac{95}{6}\,y^3x + \frac{95}{6}\,yx^3 + 40\,yx + \frac{5}{24}\,x^4 + \frac{49}{24}\,yx^4 + \frac{49}{24}\,y^4x + \frac{1}{12}\,yx^5\\
&+\frac{4}{14175}\,y^5x^5 + \frac{2}{315}\,y^5x^4 + \frac{2}{315}\,y^4x^5 + \frac{2}{45}\,y^5x^3 + \frac{46}{315}\,y^4x^4\\
&+\frac{2}{45}\,y^3x^5 + \frac{1}{9}\,y^5x^2 + \frac{47}{45}\,y^4x^3 + \frac{47}{45}\,y^3x^4 + \frac{1}{9}\,y^2x^5\\
&+\frac{1}{12}\,y^5x + \frac{8}{3}\,y^4x^2 + \frac{689}{90}\,y^3x^3 + \frac{8}{3}\,y^2x^4 + 5\,y^2\\
&+5\,x^2 + \frac{5}{3}\,y^3 + \frac{5}{3}\,x^3 + \frac{5}{24}\,y^4 + \frac{1}{120}\,y^5 + \frac{1}{120}\,x^5
\end{aligned}
\tag{4.23}
$$

- $N = 15$:

$$
\begin{aligned}
&1 + 6\,x + 80\,y^2x + 6\,y + \frac{1445}{12}\,y^2x^2 + \frac{353}{6}\,y^2x^3 + \frac{353}{6}\,y^3x^2 + 80\,yx^2\\
&+\frac{75}{2}\,y^3x + \frac{75}{2}\,yx^3 + 57\,yx + \frac{5}{8}\,x^4 + \frac{29}{4}\,yx^4 + \frac{29}{4}\,y^4x + \frac{71}{120}\,yx^5\\
&+\frac{134}{14175}\,y^5x^5 + \frac{34}{315}\,y^5x^4 + \frac{34}{315}\,y^4x^5 + \frac{1}{36}\,y^6x^2 + \frac{23}{45}\,y^5x^3\\
&+\frac{263}{210}\,y^4x^4 + \frac{23}{45}\,y^3x^5 + \frac{1}{36}\,y^2x^6 + \frac{1}{60}\,y^6x + \frac{35}{36}\,y^5x^2 + \frac{181}{30}\,y^4x^3\\
&+\frac{181}{30}\,y^3x^4 + \frac{35}{36}\,y^2x^5 + \frac{1}{60}\,yx^6 + \frac{2}{135}\,y^6x^3 + \frac{2}{135}\,y^3x^6 + \frac{1}{315}\,y^6x^4\\
&+\frac{1}{315}\,y^4x^6 + \frac{4}{14175}\,y^6x^5 + \frac{4}{467775}\,y^6x^6 + \frac{4}{14175}\,y^5x^6 + \frac{1}{720}\,x^6\\
&+\frac{1}{720}\,y^6 + \frac{71}{120}\,y^5x + \frac{187}{16}\,y^4x^2 + \frac{5339}{180}\,y^3x^3 + \frac{187}{16}\,y^2x^4 + \frac{15}{2}\,y^2\\
&+\frac{15}{2}\,x^2 + \frac{10}{3}\,y^3 + \frac{10}{3}\,x^3 + \frac{5}{8}\,y^4 + \frac{1}{20}\,y^5 + \frac{1}{20}\,x^5
\end{aligned}
\tag{4.24}
$$

- $N = 17$:

$$
\begin{aligned}
&1 + 7\,x + 7\,y + \frac{1}{945}\,y^4x^7 + 77\,yx + \frac{21}{2}\,y^2 + \frac{21}{2}\,x^2 + \frac{259}{2}\,y^2x \\
&+\frac{259}{2}\,yx^2 + \frac{931}{4}\,y^2x^2 + \frac{35}{6}\,y^3 + \frac{35}{6}\,x^3 + \frac{455}{6}\,y^3x + \frac{455}{6}\,yx^3 \\
&+\frac{567}{4}\,y^3x^2 + \frac{567}{4}\,y^2x^3 + \frac{3199}{36}\,y^3x^3 + \frac{8}{42567525}\,y^7x^7 + \frac{35}{24}\,y^4 \\
&+\frac{35}{24}\,x^4 + \frac{469}{24}\,y^4x + \frac{469}{24}\,yx^4 + \frac{1799}{48}\,y^4x^2 + \frac{1799}{48}\,y^2x^4 \\
&+\frac{1}{945}\,y^7x^4 + \frac{3457}{144}\,y^4x^3 + \frac{3457}{144}\,y^3x^4 + \frac{416}{63}\,y^4x^4 + \frac{2}{14175}\,y^5x^7 \\
&+\frac{7}{40}\,y^5 + \frac{7}{40}\,x^5 + \frac{287}{120}\,y^5x + \frac{287}{120}\,yx^5 + \frac{1121}{240}\,y^5x^2 + \frac{1121}{240}\,y^2x^5 \\
&+\frac{73}{24}\,y^5x^3 + \frac{73}{24}\,y^3x^5 + \frac{2}{14175}\,y^7x^5 + \frac{107}{126}\,y^5x^4 + \frac{107}{126}\,y^4x^5 \\
&+\frac{523}{4725}\,y^5x^5 + \frac{4}{467775}\,y^7x^6 + \frac{4}{467775}\,y^6x^7 + \frac{7}{720}\,y^6 + \frac{7}{720}\,x^6 \\
&+\frac{97}{720}\,y^6x + \frac{97}{720}\,yx^6 + \frac{4}{15}\,y^6x^2 + \frac{4}{15}\,y^2x^6 + \frac{19}{108}\,y^6x^3 + \frac{19}{108}\,y^3x^6 \\
&+\frac{47}{945}\,y^6x^4 + \frac{47}{945}\,y^4x^6 + \frac{31}{4725}\,y^6x^5 + \frac{31}{4725}\,y^5x^6 + \frac{184}{467775}\,y^6x^6 \\
&+\frac{1}{5040}\,y^7 + \frac{1}{5040}\,x^7 + \frac{1}{360}\,y^7x + \frac{1}{360}\,yx^7 + \frac{1}{180}\,y^7x^2 \\
&+\frac{1}{180}\,y^2x^7 + \frac{1}{270}\,y^7x^3 + \frac{1}{270}\,y^3x^7
\end{aligned}
\tag{4.25}
$$

- $N = 19$:

$$
\begin{aligned}
&1 + 8\,x + 8\,y + \frac{31}{1890}\,y^4x^7 + 100\,yx + 14\,y^2 + 14\,x^2 + 196\,y^2x \\
&+196\,yx^2 + \frac{819}{2}\,y^2x^2 + \frac{28}{3}\,y^3 + \frac{28}{3}\,x^3 + \frac{413}{3}\,y^3x + \frac{413}{3}\,yx^3 \\
&+\frac{896}{3}\,y^3x^2 + \frac{896}{3}\,y^2x^3 + \frac{10073}{45}\,y^3x^3 + \frac{44}{3869775}\,y^7x^7 + \frac{35}{12}\,y^4 \\
&+\frac{35}{12}\,x^4 + \frac{133}{3}\,y^4x + \frac{133}{3}\,yx^4 + \frac{1183}{12}\,y^4x^2 + \frac{1183}{12}\,y^2x^4 + \frac{31}{1890}\,y^7x^4 \\
&+\frac{1357}{18}\,y^4x^3 + \frac{1357}{18}\,y^3x^4 + \frac{104071}{4032}\,y^4x^4 + \frac{41}{14175}\,y^5x^7 + \frac{7}{15}\,y^5 \\
&+\frac{7}{15}\,x^5 + \frac{217}{30}\,y^5x + \frac{217}{30}\,yx^5 + \frac{1}{40320}\,x^8 + \frac{491}{30}\,y^5x^2 + \frac{491}{30}\,y^2x^5 \\
&+\frac{1}{2520}\,y^8x + \frac{9151}{720}\,y^5x^3 + \frac{9151}{720}\,y^3x^5 + \frac{41}{14175}\,y^7x^5 + \frac{247}{56}\,y^5x^4 \\
&+\frac{247}{56}\,y^4x^5 + \frac{3089}{4050}\,y^5x^5 + \frac{122}{467775}\,y^7x^6 + \frac{122}{467775}\,y^6x^7 + \frac{1}{1080}\,y^2x^8 \\
&+\frac{7}{180}\,y^6 + \frac{7}{180}\,x^6 + \frac{11}{18}\,y^6x + \frac{11}{18}\,yx^6 + \frac{2017}{1440}\,y^6x^2 + \frac{2017}{1440}\,y^2x^6 \\
&+\frac{1189}{1080}\,y^6x^3 + \frac{1189}{1080}\,y^3x^6 + \frac{1}{2520}\,yx^8 + \frac{2921}{7560}\,y^6x^4 + \frac{2921}{7560}\,y^4x^6 \\
&+\frac{319}{4725}\,y^6x^5 + \frac{319}{4725}\,y^5x^6 + \frac{1129}{187110}\,y^6x^6 + \frac{1}{40320}\,y^8 + \frac{1}{630}\,y^7 \\
&+\frac{1}{630}\,x^7 + \frac{127}{5040}\,y^7x + \frac{1}{1080}\,y^8x^2 + \frac{127}{5040}\,yx^7 + \frac{7}{120}\,y^7x^2 + \frac{7}{120}\,y^2x^7 \\
&+\frac{5}{108}\,y^7x^3 + \frac{5}{108}\,y^3x^7 + \frac{1}{1350}\,y^8x^3 + \frac{1}{1350}\,y^3x^8 + \frac{1}{3780}\,y^8x^4 \\
&+\frac{1}{3780}\,y^4x^8 + \frac{2}{42525}\,y^8x^5 + \frac{2}{42525}\,y^5x^8 + \frac{2}{467775}\,y^8x^6 + \frac{2}{467775}\,y^6x^8 \\
&+\frac{8}{42567525}\,y^8x^7 + \frac{8}{42567525}\,y^7x^8 + \frac{2}{638512875}\,y^8x^8
\end{aligned}
\tag{4.26}
$$

- $N = 21$: Microfossils (21).

$$
\begin{aligned}
&1 + 9\,x + 9\,y + \frac{109}{840}\,y^4x^7 + \frac{1}{8100}\,y^3x^9 + 126\,yx + 18\,y^2 + 18\,x^2 \\
&+282\,y^2x + 282\,yx^2 + \frac{1343}{2}\,y^2x^2 + 14\,y^3 + 14\,x^3 + \frac{1}{7560}\,y^9x^2 + 231\,y^3x \\
&+231\,yx^3 + \frac{2}{1403325}\,y^6x^9 + \frac{1141}{2}\,y^3x^2 + \frac{1141}{2}\,y^2x^3 + \frac{1}{85050}\,y^9x^5 \\
&+\frac{7462}{15}\,y^3x^3 + \frac{349}{1576575}\,y^7x^7 + \frac{1}{18900}\,y^9x^4 + \frac{21}{4}\,y^4 + \frac{21}{4}\,x^4 + \frac{357}{4}\,y^4x \\
&+\frac{357}{4}\,yx^4 + \frac{1}{18900}\,y^4x^9 + \frac{903}{4}\,y^4x^2 + \frac{903}{4}\,y^2x^4 + \frac{109}{840}\,y^7x^4 + \frac{4013}{20}\,y^4x^3 \\
&+\frac{4013}{20}\,y^3x^4 + \frac{184099}{2240}\,y^4x^4 + \frac{2}{638512875}\,y^8x^9 + \frac{1613}{56700}\,y^5x^7 + \frac{21}{20}\,y^5 \\
&+\frac{21}{20}\,x^5 + \frac{91}{5}\,y^5x + \frac{91}{5}\,yx^5 + \frac{1}{4480}\,x^8 + \frac{2809}{60}\,y^5x^2 + \frac{2809}{60}\,y^2x^5 \\
&+\frac{23}{5760}\,y^8x + \frac{10133}{240}\,y^5x^3 + \frac{10133}{240}\,y^3x^5 + \frac{1613}{56700}\,y^7x^5 + \frac{353047}{20160}\,y^5x^4 \\
&+\frac{353047}{20160}\,y^4x^5 + \frac{1711063}{453600}\,y^5x^5 + \frac{1061}{311850}\,y^7x^6 + \frac{1061}{311850}\,y^6x^7 + \frac{1}{8100}\,y^9x^3 \\
&+\frac{2}{189}\,y^2x^8 + \frac{7}{60}\,y^6 + \frac{7}{60}\,x^6 + \frac{41}{20}\,y^6x + \frac{41}{20}\,yx^6 + \frac{2563}{480}\,y^6x^2 + \frac{2563}{480}\,y^2x^6 \\
&+\frac{21041}{4320}\,y^6x^3 + \frac{21041}{4320}\,y^3x^6 + \frac{23}{5760}\,yx^8 + \frac{2}{1403325}\,y^9x^6 + \frac{643}{315}\,y^6x^4 \\
&+\frac{643}{315}\,y^4x^6 + \frac{50333}{113400}\,y^6x^5 + \frac{50333}{113400}\,y^5x^6 + \frac{2}{638512875}\,y^9x^8 \\
&+\frac{49279}{935550}\,y^6x^6 + \frac{1}{7560}\,y^2x^9 + \frac{1}{4480}\,y^8 + \frac{1}{140}\,y^7 + \frac{1}{140}\,x^7 + \frac{71}{560}\,y^7x \\
&+\frac{2}{189}\,y^8x^2 + \frac{71}{560}\,yx^7 + \frac{4}{42567525}\,y^7x^9 + \frac{3361}{10080}\,y^7x^2 + \frac{3361}{10080}\,y^2x^7 \\
&+\frac{221}{720}\,y^7x^3 + \frac{221}{720}\,y^3x^7 + \frac{1}{20160}\,yx^9 + \frac{1}{20160}\,y^9x + \frac{53}{5400}\,y^8x^3 \\
&+\frac{53}{5400}\,y^3x^8 + \frac{79}{18900}\,y^8x^4 + \frac{79}{18900}\,y^4x^8 + \frac{4}{97692469875}\,y^9x^9 \\
&+\frac{157}{170100}\,y^8x^5 + \frac{157}{170100}\,y^5x^8 + \frac{1}{85050}\,y^5x^9 + \frac{52}{467775}\,y^8x^6 \\
&+\frac{52}{467775}\,y^6x^8 + \frac{1}{362880}\,x^9 + \frac{62}{8513505}\,y^8x^7 + \frac{62}{8513505}\,y^7x^8 \\
&+\frac{2}{8292375}\,y^8x^8 + \frac{1}{362880}\,y^9 + \frac{4}{42567525}\,y^9x^7
\end{aligned}
\tag{4.27}
$$

Now, we apply the above confluent hypergeometric's in a sort of problems and as motivations of future works with other elliptical models and situations.

4.3.1 Biology: mouse vertebra

This problem has been studied deeply by Dryden and Mardia (1998). The data come from an investigation into the effects of selection for body weight on the shape of mouse vertebra and the experiments consider the second thoracic vertebra T2 of 30 control (C), 23 large (L) and 23 small (S) bones. The control group contains unselected mice, the large group contains mice selected at each generation according to large body weight and the small group was selected for small body weight. In order to apply the finite densities we do not consider the third landmark of the total 6, they proposed (see Dryden and Mardia (1998), p.10 and the data given in p. 313-316). One of the aims is to study configuration changes among the three groups. We suppose that normality assumption is correct according to the analysis of Dryden and Mardia (1998) in that sense.

Inference is based on (4.19), a confluent hypergeometric polynomial of degree 2 in the two eigenvalues of the zonal polynomial argument, then after a very simple computation we have the following configuration locations of the three groups.

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\sigma^2}$
Control	-0.10766	0.15599	-0.0020584	-0.97058	0.0016491
Large	-0.084652	0.12434	-0.0050449	-1.0786	0.0021276
Small	-0.091618	0.21291	-0.0062915	-1.0179	0.0019793

Then the likelihood ratios for the paired tests $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$, C-L,L-S and C-S of equal configuration locations with the corresponding p-values are given next

	Control-Large	Control-Small	Large-Small
$-2\log\Lambda \approx \chi^2_4$	45.72	26.66	50.48
p	0.0000000028	0.0000232863	0.0000000003

So, we can say that there is strong evidence for different configuration changes, and the most important is given between small and large, as we expected. These results agree with those given by Dryden and Mardia (1998), but a further comparison between the two approaches and some tests for checking uniform deformations (see step IV), can be done in a future work.

Removing a landmark at random from a figure deserves a deep study, we just explore in the next few lines some aspects in the particular context of the mouse vertebra and gorilla skulls applications, and we leave this interesting problem for a future research. In the mouse vertebra example we removed the third landmark, now we explore, a little, the importance of removing one landmark at random in this example, for this, we show in the next table, the results of all the possible removals and the corresponding tests $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of equal configuration locations with the corresponding p-values. We focus the attention on the small (S) and large groups (L); the removed landmark is specified after the group, for example, S3 refers the original group small without the third landmark; the last column of the following table gives the p-value of the S-L test.

The six anatomical landmarks on the second thoracic mouse vertebra are shown in figure 4.1. They are symmetrically selected by measuring the extreme positive and negative curvature of the bone. See Dryden and Mardia (1998) for more details.

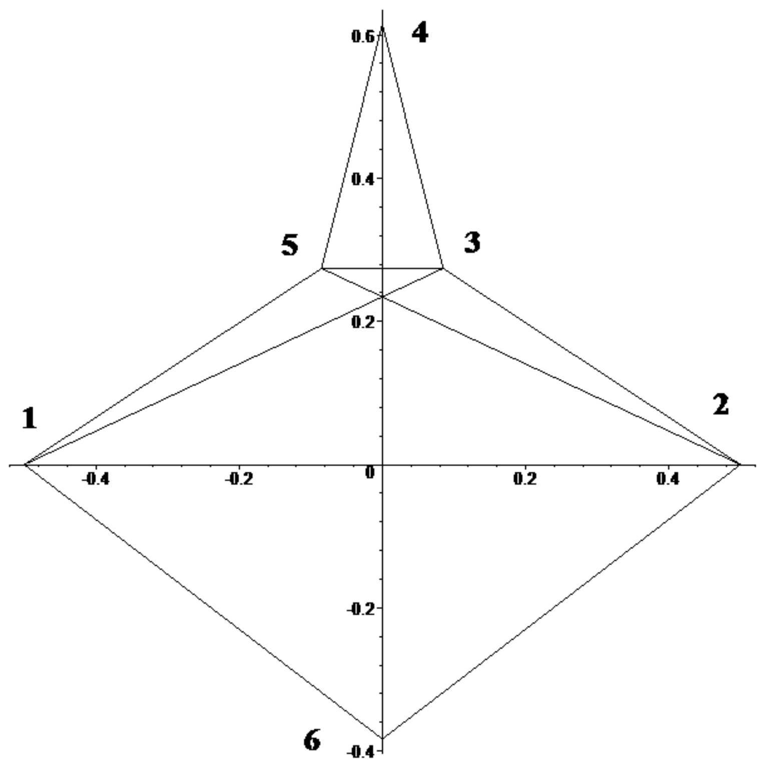

Figure 4.1: Mouse vertebra

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\sigma^2}$	p-value
S1	1.44E+00	-8.19E-01	3.18E+00	-3.50E+00	1.39E-02	6.02E-03
L1	1.42E+00	-7.43E-01	2.97E+00	-3.04E+00	1.46E-02	
S2	1.66E-01	-6.31E-02	1.19E+00	-2.30E+00	5.33E-03	1.59E-08
L2	1.90E-01	-1.35E-01	9.97E-01	-2.05E+00	5.94E-03	
S3	-9.16E-02	2.13E-01	-6.29E-03	-1.02E+00	1.98E-03	3.00E-10
L3	-8.47E-02	1.24E-01	-5.04E-03	-1.08E+00	2.13E-03	
S4	-1.90E-01	7.24E-01	1.28E-01	-1.68E+00	3.29E-03	2.10E-16
L4	-2.03E-01	7.08E-01	1.92E-01	-2.05E+00	6.70E-03	
S5	-1.81E-01	1.65E+00	1.25E-01	-1.92E+00	6.58E-03	1.82E-10
L5	-2.57E-01	1.99E+00	2.02E-01	-2.38E+00	8.73E-03	
S6	-1.82E-01	1.66E+00	-1.49E-01	3.21E-01	1.89E-03	7.70E-24
L6	-2.53E-01	2.01E+00	-1.41E-01	2.16E-01	4.41E-03	

The strong symmetry of the figure maybe explains some differences in the results, and maybe suggests that the isotropic assumption proposed by Dryden and Mardia (1998) in the original example is non appropriate. The isotropic model is not convenient because some correlations are expected between parts that lie either side of the axis of symmetry. The equality or inequality in curvature of symmetric points maybe explains the remaining differences, see the p-values of landmarks (3, 5), (4, 6) and (1, 2).

This results seem to show that the extreme values are not key points in the discrimination process. In any case we have some evidence that the two configuration populations are different independently of the landmark being removed.

Perhaps an study of this situation involving, non isotropic models, symmetry and curvature, can provide more details about the corresponding discrimination process in the mouse vertebra.

4.3.2 Biology: gorilla skulls

In this application Dryden and Mardia (1998) investigate the cranial differences between the 29 male and 30 female apes by studying 8 anatomical landmarks. The landmarks are: *pr* (1), *l* (2), *o* (3), *ba* (4), *st* (5), *na* (6), *n* (7), *b* (8).

For the finiteness of the configuration density we remove the third landmark *o* (see Dryden and Mardia (1998) p.11, and the data in p. 317-318) and the corresponding confluent hypergeometric is a polynomials of degree 4, see (4.20).

The estimators of the configuration location and scale parameters are given below

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$
Female	-0.28033	0.31315	-0.42269	-0.59672
Male	-0.33313	0.42484	-0.43594	-0.5734

$\cdots$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\sigma^2}$
$\cdots$	0.27398	-1.4695	0.7363	-1.2665	0.0042665
$\cdots$	0.30563	-1.306	0.73169	-1.0594	0.0050404

So the likelihood ratio for $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of configuration location cranial difference between the sexes of the apes, with the corresponding p-values, is the following.

	Female-Male
$-2 \log \Lambda \approx \chi^2_8$	72.94
p	0.000000000001274

This ratifies strong evidence for differences between the female and male configuration locations. These results agree with those given by Dryden and Mardia (1998), but a further comparison between the two approaches and uniform deformation tests must be done in a future work.

Now, in this case there is no symmetry, so we should expect that a random removal of one landmark does not change significantly the results of the tests. In the next table we provide all the possible removals and the corresponding tests $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of configuration location cranial difference between male (M) and female (F) groups, along with the p-values. Recall that the landmarks on the midline section of the ape cranium are: prosthion *pr* (1), lambda *l* (2), opisthion *o* (3), basion *ba* (4), staphylion *st* (5), nariale *na* (6), nasion *n* (7), bregma *b* (8). See figure 4.2. The face region is taken to be comprised of landmarks 7, 4, 5, 1 and 6. The braincase region is taken to be comprised of landmarks 7, 8, 2, 3 and 4.

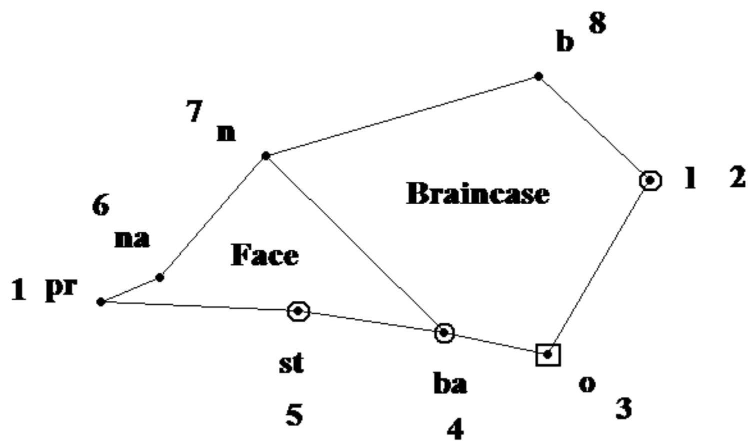

Figure 4.2: Gorilla skull

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$
M1	-1.79E+00	3.55E+00	-4.37E+00	6.22E+00	-3.58E+00	3.12E+00
F1	-1.29E+00	2.85E+00	-3.00E+00	4.81E+00	-2.98E+00	2.70E+00
M2	-6.12E-01	9.77E-01	-2.31E-01	-1.45E+00	2.41E+00	-6.29E+00
F2	-4.21E-01	4.93E-01	-8.28E-02	-2.16E+00	3.53E+00	-1.01E+01
M3	-3.33E-01	4.25E-01	-4.36E-01	-5.73E-01	3.06E-01	-1.31E+00
F3	-2.80E-01	3.13E-01	-4.23E-01	-5.97E-01	2.74E-01	-1.47E+00
M4	-4.62E-01	4.44E-01	-3.88E-01	-5.85E-01	4.82E-01	-1.33E+00
F4	-3.74E-01	2.50E-01	-3.75E-01	-5.69E-01	4.49E-01	-1.35E+00
M5	-1.70E-01	7.26E-01	-4.61E-01	-6.56E-01	4.23E-01	-1.40E+00
F5	-1.44E-01	5.98E-01	-4.35E-01	-6.59E-01	4.00E-01	-1.43E+00
M6	-1.67E-01	7.23E-01	-4.34E-01	2.71E-01	4.13E-01	-1.58E+00
F6	-1.43E-01	5.97E-01	-3.49E-01	1.02E-01	3.81E-01	-1.58E+00
M7	-1.70E-01	7.27E-01	-4.34E-01	2.72E-01	-3.81E-01	-7.28E-01
F7	-1.45E-01	5.98E-01	-3.48E-01	1.04E-01	-3.70E-01	-6.95E-01
M8	-1.68E-01	7.26E-01	-4.33E-01	2.72E-01	-3.83E-01	-7.28E-01
F8	-1.44E-01	5.98E-01	-3.49E-01	1.03E-01	-3.72E-01	-6.94E-01

Group	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\sigma^2}$	p-value
M1	-1.33E+00	-4.41E-01	2.94E-02	1.55E-15
F1	-1.50E+00	-1.84E-01	1.30E-02	
M2	3.36E+00	-6.94E+00	8.60E-02	1.78E-15
F2	4.97E+00	-1.17E+01	1.55E-01	
M3	7.32E-01	-1.06E+00	5.04E-03	1.28E-12
F3	7.36E-01	-1.27E+00	4.26E-03	
M4	8.73E-01	-1.08E+00	5.85E-03	3.90E-14
F4	8.85E-01	-1.17E+00	3.26E-03	
M5	8.23E-01	-1.13E+00	4.83E-03	3.34E-08
F5	8.45E-01	-1.23E+00	3.54E-03	
M6	8.16E-01	-1.29E+00	6.00E-03	3.59E-14
F6	8.29E-01	-1.36E+00	3.98E-03	
M7	9.54E-01	-1.44E+00	4.53E-03	1.13E-18
F7	9.60E-01	-1.52E+00	2.87E-03	
M8	4.86E-01	-1.49E+00	4.95E-03	6.35E-14
F8	4.51E-01	-1.49E+00	3.45E-03	

All the cases are very similar and support the conclusion about the difference in the configuration of both populations, and the conclusion is independent of the landmark being removed, we should expect minor extreme differences when there is not symmetry and under the isotropic model.

We noted that removing 1 landmark, the conclusion of this problem is preserved, now, the next question goes in terms of the possibility of removing three points and obtaining the same conclusion. In the following lines we removed two landmarks after removing the landmark o. We are taking advantage of our finite configuration density which can be computed easily for all possible combinations, and without approximations. In this problem of 8 points, we can remove maximum 3 landmarks in order to obtain a non trivial configuration. If the three digit after the male (M) and the female (F) groups represent the landmarks removed of the original figure, we have the following populations parameter estimations and the p-values of the corresponding tests $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of configuration location cranial difference between both groups. In this example, we remove at the third landmark (o) and all the possible groups of two points after that.

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\sigma^2}$	p-value
S312	-6.20E+00	3.56E+00	-8.29E+00	3.74E+00	7.13E-02	3.04E-12
L312	-4.02E+00	2.38E+00	-5.17E+00	2.17E+00	4.64E-02	
S314	-1.24E+00	1.25E+00	-1.32E+00	5.92E-01	6.47E-03	1.84E-17
L314	-1.02E+00	9.92E-01	-1.05E+00	3.31E-01	9.75E-03	
S315	-1.11E+00	6.72E-01	-1.20E+00	9.30E-02	4.79E-03	6.36E-11
L315	-1.02E+00	5.63E-01	-1.04E+00	1.82E-04	5.56E-03	
S316	-2.20E+00	2.01E+00	-1.60E+00	5.70E-01	9.73E-03	6.51E-10
L316	-2.23E+00	1.80E+00	-1.41E+00	3.70E-01	1.02E-02	
S317	-1.13E+00	2.11E+00	-1.88E+00	5.45E-01	1.02E-02	5.68E-18
L317	-1.37E+00	2.16E+00	-1.63E+00	2.75E-01	1.29E-02	
S318	-1.12E+00	2.11E+00	-1.98E+00	1.53E+00	1.05E-02	3.44E-14
L318	-1.38E+00	2.17E+00	-1.94E+00	1.32E+00	1.06E-02	
S324	2.43E+00	5.40E+00	3.83E+00	6.36E+00	4.07E-02	3.01E-01
L324	2.95E+00	6.05E+00	4.44E+00	7.20E+00	6.92E-02	
S325	2.53E+00	5.39E+00	3.57E+00	6.22E+00	4.89E-02	3.20E-02
L325	3.08E+00	6.25E+00	4.07E+00	6.99E+00	7.89E-02	
S326	9.69E-01	-1.30E+01	2.10E+00	-1.47E+01	3.43E-01	4.88E-07
L326	3.67E-01	-6.57E+00	1.13E+00	-7.19E+00	1.02E-01	
S327	-4.53E-01	-3.00E+00	2.76E+00	-2.21E+01	2.39E-01	1.74E-13
L327	-5.05E-01	-1.27E+00	1.36E+00	-9.29E+00	6.60E-02	
S328	-4.29E-01	-3.06E+00	1.47E+00	-1.67E+01	2.19E-01	1.07E-14
L328	-5.08E-01	-1.24E+00	5.47E-01	-6.88E+00	6.91E-02	

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\sigma^2}$	p-value
S345	2.96E+00	6.18E+00	2.91E+00	4.90E+00	9.89E-02	2.80E-03
L345	3.24E+00	6.44E+00	2.93E+00	4.58E+00	2.01E-01	
S346	-3.35E-01	-2.03E+00	2.21E-01	-1.62E+00	1.14E-02	6.18E-17
L346	-2.46E-01	-1.53E+00	2.88E-01	-1.14E+00	1.38E-02	
S347	-5.37E-01	-5.80E-01	2.68E-01	-2.01E+00	7.54E-03	2.08E-23
L347	-5.42E-01	-4.76E-01	3.59E-01	-1.43E+00	9.65E-03	
S348	-5.38E-01	-5.76E-01	-2.08E-01	-1.96E+00	8.92E-03	6.37E-16
L348	-5.43E-01	-4.74E-01	-1.18E-01	-1.48E+00	1.07E-02	
S356	1.09E-01	-1.43E+00	5.71E-01	-1.14E+00	6.80E-03	1.89E-08
L356	1.25E-01	-1.25E+00	5.52E-01	-9.25E-01	6.77E-03	
S357	-4.74E-01	-5.03E-01	7.24E-01	-1.39E+00	4.34E-03	8.14E-13
L357	-4.98E-01	-4.50E-01	7.14E-01	-1.14E+00	4.31E-03	
S358	-4.76E-01	-5.01E-01	2.37E-01	-1.36E+00	5.51E-03	7.98E-07
L358	-5.00E-01	-4.50E-01	2.59E-01	-1.18E+00	4.93E-03	
S367	-2.80E-01	3.16E-01	6.74E-01	-1.60E+00	5.21E-03	1.14E-17
L367	-3.33E-01	4.27E-01	6.71E-01	-1.36E+00	6.46E-03	
S368	-2.80E-01	3.15E-01	1.86E-01	-1.56E+00	6.46E-03	2.05E-11
L368	-3.32E-01	4.27E-01	2.14E-01	-1.40E+00	6.82E-03	
S378	-2.79E-01	3.18E-01	-4.14E-01	-6.07E-01	2.36E-03	6.61E-14
L378	-3.30E-01	4.27E-01	-4.27E-01	-5.81E-01	2.42E-03	

Note that even in the case where the first removal is at random, in this case the landmark o, the second and third removals preserve the conclusion, in 18 times of the possible 21 combinations, about rejecting the null hypothesis, i.e. the strong evidence for differences in both populations is preserved. However, the opposite results of combinations 324 and 325, and 345 in a minor sense, maybe is explained by the approximately collinearity of points 3, 4, 5 and the importance of this three combinations for comprising the braincase and the curvature information (see Dryden and Mardia (1998)). But it deserves a deep study in this particular combinations.

In any case one random landmark removal do not alter sufficiently the results of the mouse and gorilla experiments, even in symmetric figures, but 3 or more removals requires additional studies.

However we can improve our conclusions and avoid the false positive about rejecting the null hypothesis by using a selecting model criterion in this two particular examples.

Explicitly, we can apply the Schwarz's criterion in order to select the best candidate to be removed, for example, landmark 6 and landmark n are maybe the best selections in the mouse and gorilla skulls problems, respectively, see the next table. The Schwarz criterion is based on the configuration likelihood instead of the classical elliptical likelihood, because we can take advantage of the finite configuration, which in this two examples is a polynomial of degree two and five respectively.

M-F landmark removed	Schwarz's crit.	S-L landmark removed	Schwarz's crit.
pr	8.74E+01	1	1.44E+01
l	8.71E+01	2	4.21E+01
o	7.29E+01	3	5.05E+01
ba	8.05E+01	4	7.96E+01
st	5.05E+01	5	5.14E+01
na	8.07E+01	6	1.15E+02
n	1.03E+02		
b	7.94E+01		

In any case, we have noticed that in the mouse vertebra and gorilla skull applications, the conclusions do not differ to much from the tests based on the landmarks we removed (3 and o, respectivelly) in our examples, and the equality of both configuration mean populations, in any case, is highly rejected.

As a conclusion of this exploratory examples, we note that symmetry (joint with curvature) and the isotropic model play an important role in the landmark removal procedure and it is important to potentiate the use of our finite configuration densities in this two situations.

4.3.3 Biology: The university school study subsample

In this experiment Bookstein (1991) studies sex shape differences between 8 craniofacial landmarks for 36 normal Ann Arbor boys and 26 girls near the ages of 8 years. In order to get a finite configuration density we discard the landmark Sella (see Bookstein (1991), p.401-405), then the hypergeometric functions is a polynomial of degree 4, see (4.20). We can study all the possible combinations of landmarks as in the above example, we have similar conclusions, because the figures are not symmetric.

Then, the estimators of the configuration location and scale parameters are the following

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$
Male	-1.2425	2.1948	0.46435	-1.3752
Female	-1.2483	2.2331	0.43685	-1.3845

⋯	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\sigma^2}$
⋯	-0.91487	0.66127	0.15775	-0.069042	0.0032908
⋯	-0.92903	0.70439	0.1616	-0.077236	0.0059142

In this case, the likelihood ratio for $H_0 : \mathcal{U}_1 = \mathcal{U}_2$ vs $H_a : \mathcal{U}_1 \neq \mathcal{U}_2$ of configuration location cranialfacial difference between boys and girls, with the corresponding p-values, is the following.

	Male-Female
$-2\log\Lambda \approx \chi^2_8$	5.48
p	0.7053

And the difference between these two configuration locations is insignificant. A similar global conclusion gives Bookstein (1991), however a more detailed study of landmark subsets is required, then possible differences can be detected, as Bookstein (1991) ratifies in a different shape context.

4.3.4 Medicine: brain MR scans of schizophrenic patients

We return to the applications in Dryden and Mardia (1998), in this case, they study 13 landmarks on a near midsagittal two dimensional slices from magnetic resonance (MR) brain scans of 14 schizophrenic patients and 14 normal patients. Given that the number of two dimensional landmarks is odd we preserve them leading a 10 degree confluent hypergeometric polynomial, easy to compute, see (4.23).

Thus, the estimators of the configuration location and scale parameters are given by

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$
Normal	-0.64099	2.6942	-1.2744	-2.8323	-0.42155	-1.003
Squizo.	-0.68623	2.393	-1.145	-2.8484	-0.37349	-1.0744

Group	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\mathcal{V}_{51}}$	$\widetilde{\mathcal{V}_{52}}$	$\widetilde{\mathcal{V}_{61}}$	$\widetilde{\mathcal{V}_{62}}$
Normal	-0.31011	-2.3094	-0.30236	-3.5261	0.36	-0.90135
Squizo.	-0.23173	-2.1929	-0.20173	-3.3226	0.38123	-0.84316

Group	$\widetilde{\mathcal{V}_{71}}$	$\widetilde{\mathcal{V}_{72}}$	$\widetilde{\mathcal{V}_{81}}$	$\widetilde{\mathcal{V}_{82}}$	$\widetilde{\mathcal{V}_{91}}$	$\widetilde{\mathcal{V}_{92}}$
Normal	0.1597	-2.2205	0.8518	-0.7578	1.8686	0.86501
Squizo.	0.20429	-2.109	0.84683	-0.56588	1.7948	0.88466

Group	$\widetilde{\mathcal{V}_{10,1}}$	$\widetilde{\mathcal{V}_{10,2}}$	$\widetilde{\sigma^2}$
Normal	-0.14205	0.20718	0.010843
Squizo.	-0.079005	0.1378	0.054064

Dryden and Mardia (1998) advert about the small sample size of this experiment and obviously this can explain the opposite result

	Normal-Squizophrenic
$-2\log\Lambda \approx \chi^2_{20}$	11.96
p	0.9174

Mean shape difference is concluded in Dryden and Mardia (1998), but configuration difference is definitely insignificant. The controversial configuration location results could suggest a deep study for small sample likelihood and perhaps it can ratify important different conclusions of studies about schizophrenia classification based only on MR scans. But the most important fact here is the geometric meaning of the data, because it certainly differs from the preceding applications, which have an explicit geometric explanation.

4.3.5 Image analysis: postcode recognition

Again, a 13 landmark problem, which supposes a 10 degree confluent hypergeometric appears, in this case Dryden and Mardia (1998) studies a 30 random sample of handwritten digit 3 for postcode recognition. The data is available in Dryden and Mardia (1998), p. 318-320.

The next table shows, the configuration location and scale parameter estimates, joint the configuration coordinates of a template number 3 digit, with two equal sized arcs, and 13 landmarks (two coincident) lying on two regular octagons see Dryden and Mardia (1998), p.153.

Group	$\widetilde{\mathcal{V}_{11}}$	$\widetilde{\mathcal{V}_{12}}$	$\widetilde{\mathcal{V}_{21}}$	$\widetilde{\mathcal{V}_{22}}$	$\widetilde{\mathcal{V}_{31}}$	$\widetilde{\mathcal{V}_{32}}$	$\widetilde{\mathcal{V}_{41}}$	$\widetilde{\mathcal{V}_{42}}$	$\widetilde{\mathcal{V}_{51}}$	$\widetilde{\mathcal{V}_{52}}$	$\widetilde{\mathcal{V}_{61}}$	$\widetilde{\mathcal{V}_{62}}$	$\widetilde{\mathcal{V}_{71}}$	$\widetilde{\mathcal{V}_{72}}$	$\widetilde{\mathcal{V}_{81}}$	$\widetilde{\mathcal{V}_{82}}$	$\widetilde{\mathcal{V}_{91}}$	$\widetilde{\mathcal{V}_{92}}$
Digit 3	-0.79087	1.9432	-2.1073	1.5875	-2.713	0.81862	-2.8084	-0.066901	-2.5712	0.71315	-2.6934	1.2955	-3.1548	1.6802	-3.8004	1.34	-4.0517	0.33141
Template	-2.0908	2.2071	-4.0409	2.8051	-4.5904	2.2904	-4.2069	1.3688	-3.3126	1.7582	-3.5881	2.7053	-5.4996	4.0629	-7.5557	4.8428	-8.2514	4.4208

$\widetilde{\mathcal{V}_{10,1}}$	$\widetilde{\mathcal{V}_{10,2}}$	$\widetilde{\sigma^2}$
-3.7659	-0.6583	0.22904
-6.9108	2.8899	

And clearly, this enormous difference must be revealed in the corresponding test

	Digit-Template
$-2\log\Lambda \approx \chi^2_{20}$	494.88
p	≈ 0

This result was corroborated with probability ≈ 0.0002 by Dryden and Mardia (1998), p. 153. under a shape model. In any case there is strong evidence that the configuration location does not have the configuration of the ideal template for digit 3.

Remark 4.3.1. Finally, we must note that the remaining bidimensional applications in Dryden and Mardia (1998), and Bookstein (1991), etc. can be studied with the finite configuration densities and exact formulae for zonal polynomials; in fact the three dimensional applications available in the literature (see Goodall and Mardia (1993)) and others in genetics for 3D DNA part, etc, can be studied in an exact form with the help of corollary 4.1.5 via lemma 4.2.1 and exact formulae for zonal polynomials of third degree in James (1964), avoiding the open truncation problems implicit in Koev and Edelman (2006).

Of course the study of finite configuration densities associated to Pearson, Bessel, Jensen-logistic and the general Kotz, will facilitate exact inference and will avoid the addressed truncation problem, but it will depends on some developments in integration and series representation. This topic is been investigated.

Conclusions

- This thesis provides the necessary mathematical tools in integration and partition theory, for deriving the noncentral configuration density of any elliptical model by the computation of a simple single integral.

- It avoids the multivariate calculation performed in the published works, the isotropic Gaussian and some supported Pearson VII type integrals, both of them here revised and corrected.

- Exact expressions for the classical elliptical families are derived by using some partition results and single integration, explicitly:

  - Kotz configuration density,
  - Pearson VII type configuration density,
  - Bessel configuration density,
  - Jensen-Logistic configuration density.

- The general procedure for performing inference of any elliptical model is proposed and it is set in such manner that the published efficient numerical algorithms for confluent infinite series type involving zonal polynomials, can be used.

- Moreover, a further simplification of the closed computational problem is also proposed: the study of finite configuration densities.

- Then a subfamily of finite configurations is derived and as a simple example of their use, exact inference for testing configuration location differences in some applied problems is provided.

- Thus, by using our formulae for zonal polynomials, some two dimensional applications of the shape literature are studied.

- The applications include:

  - Biology: mouse vertebra,
  - Biology: gorilla skulls,
  - Biology: girl and boy craniofacial studies,
  - Medicine: brain MR scans of schizophrenic patients,
  - Image analysis: postcode recognition.

- Given the simplicity in computations involving finite configuration densities, we can study deeply the landmark selection methodology surveyed in the mouse vertebra and gorilla skulls applications; they suggest to study the symmetry (joint with curvature) and the isotropy of the model in order to characterize the sensibility of the classification and the construction of model criteria for this particular and exceptional finite configurations.

- Some of the theoretical results of the thesis are summarized in Caro-Lopera *et al* (2008a).

- The inference and applications of the thesis are surveyed in Caro-Lopera *et al* (2008b).

Future works

We can divide the future research in two groups:

Research area I.

- Matrix Generalized Kummer relations.
- Finite noncentral elliptical configuration densities.
- Applications with exact densities avoiding the truncation problem of Koev and Edelman (2006).

Research area II.

- Construction of new shape models via hypergeometric induction.
- Matrix Generalized Euler relations.
- Finite shape densities via Euler relations and applications.

Bibliography

F. L. Bookstein. (1991). *Morphometric tools for landmark data*. Cambridge, England: Cambridge University Press.

F. J. Caro Lopera and D. K. Nagar. (2006). Evaluation of matrix Liouville-Dirichlet integrals using Laplace transform. *Integral Transforms and Special Functions*, 17 (4), 245-255.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2006). A Formula for Complex Zonal Polynomials of Second Order. *Revista de Matemática: Teoría y Aplicaciones*, 13 (1), 35–39.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2007). A formula for Jack polynomials of the second order. *Applicationes Mathematicae* (Warsaw) 34, 113-119.

F. J. Caro-Lopera, G. González-Farías and N. Balakrishnan. (2008). *Partitions and some related aspects via C-arrays.* 223 p. Project book at Birkhauser, to be submitted.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2008a). Noncentral Elliptical Configuration Density. Submitted.

F. J. Caro-Lopera, J. A. Díaz-García, and G. González-Farías. (2008b). Inference in statistical shape theory: Elliptical Configuration Densities. Submitted.

A.G. Constantine. (1963). Noncentral distribution problems in multivariate analysis. *The Annals of Mathematical Statistics*, 34,1270–1285.

A. W. Davis. (1980). Invariant polynomials with two matrix arguments, extending the zonal polynomials. *Multivariate analysis, V* (P.R. Krishnaiah, Ed.), 287–299, North-Holland, Amsterdam-New York.

J. A. Díaz-García, J. R. Gutiérrez, and R. Ramos. (2003). Size-and-Shape Cone, Shape Disk and Configuration Densities for the Elliptical Models. *Brazilian Journal of Probability and Statistics*, 17, 135-146.

J. A. Díaz-García, and R. Gutiérrez-Jáimez (2006). Wishart and Pseudo-Wishart Distributions under Elliptical Laws and related distributions in the Shape Theory Context. *Journal of Statistical Planning and Inference*, 136(12), 4176-4193.

J. A. Díaz-García, and F. J. Caro-Lopera. (2006). An Alternative Approach for Deriving the Laplace-Beltrami Operator for the Zonal Polynomials of Positive Semidefinite and Definite Matrix Argument. *Far East Journal of Mathematical Sciences*, 22 (3), 273–290.

I. L. Dryden and K.V. Mardia. (1998). *Statistical shape analysis.* John Wiley and Sons, Chichester.

I. Dumitriu, A. Edelman, and G. Shuman. (2004). MOPS:Multivariate orthogonal polynomials (simbolically). *http://math.berkeley.edu/ dumitriu/mopspage.html*

K.T. Fang, and Y. T. Zhang. (1990). *Generalized Multivariate Analysis*. Science Press, Springer-Verlag, Beijing.

K.T. Fang, S. Kotz, and K.W.Ng. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, New York.

C. R. Goodall, and K.V. Mardia. (1993). Multivariate Aspects of Shape Theory. *The Annals of Statistics*, 21, 848–866.

A. K. Gupta, and T. Varga. (1993). *Elliptically Contoured Models in Statistics*, Kluwer Academic Publishers, Dordrecht.

R. Gutiérrez, J. Rodríguez, and A. J. Sáez. (2000). Approximation of Hypergeometric Functions with Matricial Argument through their Development in Series of Zonal Polynomials. *Electronic Transactions on Numerical Analysis*, 11, 121–130.

C. S. Herz. (1955). Bessel Functions of Matrix Argument. *The Annals of Mathematical Statistics*,61, 474–523.

A. T. James. (1964). Distributions of matrix variate and latent roots derived from normal samples, *The Annals of Mathematical Statistics*, 35, 475–501.

A. T. James. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *The Annals of Mathematical Statistics*, 39, 1711–1718.

D. R. Jensen. (1985). Multivariate distributions. *Encyclopedia of Statistical Sciences*, 6, (eds S., Kotz, N. L., Johnson, and C. B. Read), Wiley, 43–55.

D. G. Kendall, D. Barden, T. K. Carne, and H. Le. (1999). *Shape and Shape Theory.* Wiley Series in Probability and Statistics. John Wiley & Sons, LTD., Chichester.

C. G. Khatri. (1966). On certain distribution problems based on positive definite quadratic functions in normal vectors. *The Annals of Mathematical Statistics*, 37, 468–479.

P. Koev and A. Edelman. (2006). The efficient evaluation of the hypergeometric function of a matrix argument. *Mathematics of Computation* 75, 833-846.

H. Le, and D. G. Kendall. (1993). The riemannian structure of euclidean shape space: A novel environment for statistics. *The Annals of Statistics* 21, 1225-1271.

D. Littlewood. (1950). *The theory of group characters.* Oxford, Clarendon press.

K.V. Mardia, and I. L. Dryden. (1989). The Statistical Analysis of Shape Data. *Biometrika*, 76, 271–281.

R. J. Muirhead. (1982). *Aspects of multivariate statistical theory.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

L. Runze. (1997). The expected values of invariant polynomials with matrix argument of elliptical distributions. *Acta Mathematicae Applicatae Sinica*, 13 (1), 64–70.

P. Sawyer. (1997). Spherical Functions on Symmetric Cones, *Transactions of the American Mathematical Society.*, 349, 3569–3584.

G. Schwarz. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6 (2), 461–464.

C. G. Small. (1996). *The Statistical Theory of Shape.* Springer-Verlag, New York.

C. Teng, H. Fang, and W. Deng. (1989). The generalized noncentral Wishart distribution. *J. of Mathematical Research and Exposition*, 9, 479-488.

J.L. Xu, and K.T. Fang. (1989). The expected values of zonal polynomials of elliptically contoured distributions. In *Statistical Inference in Elliptically Contoured and Related Distributions* (K. T. Fang and T.W. Anderson, ed.), 469–479., Allerton Press Inc., New York.