

ZONAL POLYNOMIALS OF POSITIVE SEMIDEFINITE MATRIX ARGUMENT

José A. Díaz-Gracia and Francisco José Caro

Comunicación Técnica No I-04-06/05-07-2004
(PE/CIMAT)



ZONAL POLYNOMIALS OF POSITIVE SEMIDEFINITE MATRIX ARGUMENT

BY JOSÉ A. DÍAZ-GARCÍA

Universidad Autónoma Agraria Antonio Narro

AND

FRANCISCO JOSÉ CARO

Universidad de Antioquia

By using the linear structure theory of Magnus (12), this work proposes an alternative way to James (11) for obtaining the Laplace-Beltrami operator, who has the zonal polynomials of positive definite matrix argument as eigenfunctions, in particular, an explicit expression for the matrix $G(\mathbf{v}(X))$, which appears in the metric differential form $(ds)^2 = d\mathbf{v}'(X)G(\mathbf{v}(X))d\mathbf{v}(X)$, is obtained; also, the invariance of $(ds)^2$ under congruence transformations is proved. Explicit forms for $(ds)^2$ and $G(\mathbf{v}(X))$ are also shown under the spectral decomposition $X = HYH'$. In a new approach -apart from the classical theory of James (11)- a differential metric depending on the Moore-Penrose inverse is proposed for the space of $m \times m$ positive semidefinite matrices. As in the definite case, the Laplace-Beltrami operator for the calculation of zonal polynomials of positive semidefinite matrix argument is derived. In a parallel way the invariance of $(ds)^2$ is shown and explicit expressions for the metric and the matrix $G(\cdot)$ are obtained in terms of X and its spectral decomposition. Finally, an efficient computational method for calculating the zonal polynomials of positive semidefinite matrices are presented.

AMS 2000 Subject Classification: Primary 62H05
Secondary 33A70

KEY WORD AND PHRASES: Laplace-Beltrami operator, linear structures, zonal polynomials of positive semidefinite matrix argument

1. Introduction. It is known that zonal polynomials are defined for symmetric matrices, not necessarily positive definite (see James (8)). Nevertheless, classical problems in multivariate analysis which use zonal polynomials -noncentral distributions, computations of some matrix moments, etc.- have used non-singular random matrices. Plenty of properties and computational methods for zonal polynomials have been developed for positive definite matrices, see James (7) and (9), Constantine (1) and (2), Muirhead (14), Takemura (17), Farrell (5), etc. Recently, some works have appeared extending these problems to the case of singular random matrices, which require the computation of zonal polynomials of positive semidefinite matrix argument, see Díaz-García, et al. (3) and Díaz-García and González-Farías (4).

Several methods for computing the zonal polynomials of positive definite matrices have been proposed, but we can recognize in James (11) (proved completely by Muirhead (14)) the best computationally speaking; an algorithm for generating the coefficients can be found

in McLaren (13). Nevertheless, the construction and computation of zonal polynomials of positive semidefinite matrix argument have been rarely mentioned in the literature, v.g. see page 227 in Muirhead (14).

Here we follow the idea of James (11), namely, the zonal polynomials are zonal spherical functions in terms of the theory of symmetric spaces of Helgason (6), thus these polynomials are eigenfunctions of the Laplace-Beltrami operator, (see Helgason (6) eq. (4), p.387). In this work we find the polynomials for positive definite matrices by the linear structure theory proposed by Magnus (12), the highlights of James (11) are obtained by using the new approach and are presented in Section 3; they are: the metric $(ds)^2$ and its invariance under congruence transformations; the matrix $G(\cdot)$ which appears in $(ds)^2 = d\mathbf{v}'(X)G(\mathbf{v}(X))d\mathbf{v}(X)$; the expressions for $(ds)^2$, $G(\cdot)$ and $\det(G(\cdot))$ under the spectral decomposition of X , $X = HYH'$. In Section 4, it is proposed a differential metric for the space of the $m \times m$ positive semidefinite matrices, using the generalized inverse of Moore-Penrose, which will be invariant under congruence transformations LXL' , when L is orthogonal. Distinguishing the mathematically independent elements in the positive semidefinite matrix X , and based again in the linear structure theory, we give the expressions for the metric $(ds)^2$ and the matrix $G(\cdot)$; nevertheless, given that the base (matrix) for the structure is not unique, it is not possible to study the problem analogously to the case of positive definite matrices; but it can be done by working under the non-singular part of the spectral decomposition $X = HYH'$. That section is concluded constructing the partial differential equations for the zonal polynomials in the semidefinite case. In sections 5 and 6 we give the recurrence relation and the computation of coefficients for the zonal polynomials of positive semidefinite matrix argument, respectively. As in the definite case, the method exposed can be considered the best computationally speaking.

Now a question arises: Is it possible to define a differential metric in the space of $m \times m$ positive semidefinite matrices in terms of another kind of generalized inverse? The answer will be supported in Section 7 and we will see that it is not possible, excepting in the case treated in Section 3.

2. The Laplace-Beltrami Operator. First we give a summary of some differential geometry concepts taken from Helgason (6). Let M be a \mathbb{C}^∞ manifold. A *Pseudo-Riemannian structure* on M is a tensor field g of type $(0, 2)$, namely is an element of $\mathcal{D}_2^0(M) = \mathcal{D}_2(M)$ which is contravariant of degree 0 and covariant of degree 2 -see Helgason (6), Chapter 1, section 2- and satisfies $g(X, Y) = g(Y, X)$ for all $X, Y \in \mathcal{D}^1(M)$ -the set of all vector fields on M - and g_p is a nondegenerate bilinear form on $M_p \times M_p$, for each $p \in M$. A *Pseudo-Riemannian manifold* is a connected \mathbb{C}^∞ manifold with a pseudo-Riemannian structure. When g_p is positive definite for each $p \in M$, we say a Riemannian structure and Riemannian manifold.

Now, a pseudo-Riemannian manifold M always possesses a differential operator, one of them is called the *Laplace-Beltrami Operator*, which we define following Helgason (6), chapter X, section 2: Let g denote the pseudo-Riemannian structure on M and let $\varphi : q \rightarrow (x_1(q), \dots, x_m(q))$ be a coordinate system valid on an open set $U \subset M$, define the functions g_{ij} , g^{ij} , \tilde{g} on U by

$$g_{ij} = g \left(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j} \right),$$

$$\sum_j g_{ij} g^{jk} = \delta_{ik},$$

$$\tilde{g} = |\det(g_{ij})|.$$

Each function $f \in \mathbb{C}^\infty(M)$ gives rise to a vector field gradient of f on M whose restriction to U is given by

$$\text{grad } f = \sum_{i,j} g^{ij} \frac{\partial f}{\partial x_i} \frac{\partial}{\partial x_j}.$$

Also, if $X = \sum X_i \frac{\partial}{\partial x_i}$ on U , is a vector field on M , the divergence of X is the function on M which on U is given by

$$\text{div } X = \frac{1}{\sqrt{\tilde{g}}} \sum_i \frac{\partial}{\partial x_i} \left(\sqrt{\tilde{g}} X_i \right).$$

where the right hand side of the last two expressions can be shown to be invariant under coordinate system changes.

Thus the Laplace-Beltrami operator Δ is defined by

$$\Delta f = \text{div grad } f,$$

for $f \in \mathbb{C}^\infty(M)$.

In terms of local coordinates we have

$$\Delta f = \frac{1}{\sqrt{\tilde{g}}} \sum_k \partial_k \left(\sum_i g^{ik} \sqrt{\tilde{g}} \partial_i f \right), \quad (1)$$

which shows that Δ is a differential operator on M .

3. Zonal Polynomials of positive definite matrix argument: An approach from Linear Structures. Before considering the semidefinite case we give a different approach to James (1968) and Muirhead (14) in constructing the zonal polynomials from linear structure theory (Magnus (12)). Let $X \in S_m$ an $m \times m$ positive definite matrix, the metric differential for S_m is defined to be

$$(ds)^2 = \text{tr}(X^{-1} dX X^{-1} dX),$$

$dX = (dx_{ij})$, which is invariant under the congruent transformation

$$X \rightarrow LXL', \quad (2)$$

where $L \in Gl(m, \mathbb{R})$ is the group of $m \times m$ nonsingular real matrices..

Now,

$$\text{tr}(BX'CXD) = \text{vec}' X (B'D' \otimes C) \text{vec } X = \text{vec}' X (DB \otimes C') \text{vec } X$$

Thus

$$\begin{aligned}
(ds)^2 &= \text{vec}' dX (X^{-1} \otimes X^{-1}) \text{vec} dX \\
&= d \text{vec}' X (X^{-1} \otimes X^{-1}) d \text{vec} X \\
&= d \mathbf{v}'(X) D'_m (X^{-1} \otimes X^{-1}) D_m d \mathbf{v}(X)
\end{aligned}$$

where $d \text{vec} X = D_m \mathbf{v}(X)$, with D_m the duplication matrix of X symmetric and $\mathbf{v}(X)$ is the vectorization of X including only the mathematically independent elements of X (see equation 4.14, p. 55 in Magnus (12)). Thus the metric has the form

$$(ds)^2 = d \mathbf{v}'(X) G(\mathbf{v}(X)) d \mathbf{v}(X),$$

with

$$G(\mathbf{v}(X)) = D'_m (X^{-1} \otimes X^{-1}) D_m$$

Then

$$\det G(\mathbf{v}(X)) = 2^{m(m-1)/2} (\det X)^{-(m+1)}$$

and

$$G(\mathbf{v}(X))^{-1} = D_m^+ (X \otimes X) D_m^+,$$

see Theorem 4.11, p. 65 in Magnus (12).

Thus the operator Δ_X^* of Muirhead (14) can be written as

$$\begin{aligned}
\Delta_X^* &= (\det G(\mathbf{v}(X)))^{-1/2} \frac{\partial'}{\partial \mathbf{v}(X)} \left[(\det G(\mathbf{v}(X)))^{1/2} G(\mathbf{v}(X))^{-1} \frac{\partial}{\partial \mathbf{v}(X)} \right] \\
&= (\det X)^{(m+1)/2} \frac{\partial'}{\partial \mathbf{v}(X)} (\det X)^{-(m+1)/2} D_m^+ (X \otimes X) D_m^+ \frac{\partial}{\partial \mathbf{v}(X)},
\end{aligned}$$

note that $\mathbf{v}(X)$ is \mathbf{x} in Muirhead (14).

Take $Z = LXL'$, let us see that $\Delta_X^* = \Delta_Z^*$, i.e. Δ_X^* is invariant under congruent transformations.

Then $\text{vec} Z = (L \otimes L) \text{vec} X$ and $D_m \mathbf{v}(Z) = (L \otimes L) D_m \mathbf{v}(X)$. Thus

$$\mathbf{v}(Z) = D_m^+ (L \otimes L) D_m \mathbf{v}(X),$$

because $D_m^+ D_m = I_{m(m+1)/2}$ (where $D_m^+ (L \otimes L) D_m = T_L$ and $\mathbf{v}(Z) = \mathbf{z}$ in Muirhead (14), equation (26), p-240). Then

$$d \mathbf{v}(Z) = D_m^+ (L \otimes L) D_m d \mathbf{v}(X)$$

and

$$\frac{\partial}{\partial \mathbf{v}(X)} = (D_m^+ (L \otimes L) D_m)' \frac{\partial}{\partial \mathbf{v}(Z)}.$$

So

$$\frac{\partial}{\partial \mathbf{v}(Z)} = (D_m^+ (L \otimes L) D_m)^{-1} \frac{\partial}{\partial \mathbf{v}(X)}$$

Replacing in the metric form, we get

$$\begin{aligned}(ds)^2 &= d\mathbf{v}'(X)G(X)d\mathbf{v}(X) = d\mathbf{v}'(Z)G(Z)d\mathbf{v}(Z) \\ &= d\mathbf{v}'(X)D'_m(L' \otimes L')D_m^+ D'_m ((LXL')^{-1} \otimes (LXL')^{-1}) D_m D_m^+(L \otimes L)D_m d\mathbf{v}(X)\end{aligned}$$

Given that $(ds)^2$ is invariant under the transformation $X \rightarrow LXL' = Z$ and $d\mathbf{v}(Z) = D_m^+(L \otimes L)D_m d\mathbf{v}(X)$, then

$$G(\mathbf{v}(X)) = D'_m(L' \otimes L')D_m^+ D'_m \left((L \otimes L)^{-1}(X^{-1} \otimes X^{-1})(L' \otimes L)^{-1} \right) D_m D_m^+(L \otimes L)D_m$$

or alternatively

$$G(D_m^+(L \otimes L)D_m \mathbf{v}(X)) = (D'_m(L' \otimes L')D_m^+)^{-1}(D'_m(X^{-1} \otimes X^{-1})D_m)(D_m^+(L \otimes L)D_m)^{-1}$$

But $(D_m^+(L \otimes L)D_m)^{-1} = D_m^+(L^{-1} \otimes L^{-1})D_m$ and $D_m D_m^+ = I$. Thus

$$\begin{aligned}G(D_m^+(L \otimes L)D_m \mathbf{v}(X)) &= D'_m \left(L'^{-1} \otimes L'^{-1} \right) (X^{-1} \otimes X^{-1}) (L^{-1} \otimes L^{-1}) D_m \\ &= D'_m \left((LXL')^{-1} \otimes (LXL')^{-1} \right) D_m\end{aligned}$$

That is, under the transformation $X \rightarrow LXL'$, the matrix $G(\mathbf{v}(X))$ is transformed in

$$G(\mathbf{v}(X)) \rightarrow D'_m \left((LXL')^{-1} \otimes (LXL')^{-1} \right) D_m$$

In this way

$$\begin{aligned}\Delta_Z^* &= \Delta_{LXL'}^* \\ &= (\det G(\mathbf{v}(Z)))^{-1/2} \frac{\partial'}{\partial \mathbf{v}(Z)} \left[(\det G(\mathbf{v}(Z)))^{1/2} G(\mathbf{v}(Z))^{-1} \frac{\partial}{\partial \mathbf{v}(Z)} \right] \\ &= (\det G(D_m^+(L \otimes L)D_m \mathbf{v}(X)))^{-1/2} \frac{\partial'}{\partial \mathbf{v}(Z)} \left[(\det G(D_m^+(L \otimes L)D_m \mathbf{v}(X)))^{1/2} \right. \\ &\quad \left. (D_m^+(L^{-1} \otimes L^{-1})D_m) D_m^+(LXL' \otimes LXL') D_m^+ \right. \\ &\quad \left. D'_m \left((LXL')^{-1} \otimes (LXL')^{-1} \right) D_m^+ \frac{\partial}{\partial \mathbf{v}(X)} \right] \\ &= (\det LXL')^{(m+1)/2} \frac{\partial'}{\partial v(X)} \left[(\det LXL')^{-(m+1)/2} D_m^+(L^{-1} \otimes L^{-1}) \right. \\ &\quad \left. (LXL' \otimes LXL') \left(L'^{-1} \otimes L'^{-1} \right) D_m^+ \frac{\partial}{\partial v(X)} \right] \\ &= (\det L)^{m+1} (\det X)^{(m+1)/2} \frac{\partial'}{\partial v(X)} \left[(\det L)^{-(m+1)} (\det X)^{-(m+1)/2} \right. \\ &\quad \left. (D_m^+(X \otimes X) D_m^+ \frac{\partial}{\partial v(X)}) \right] \\ &= (\det X)^{(m+1)/2} \frac{\partial'}{\partial v(X)} \left[(\det X)^{-(m+1)/2} (D_m^+(X \otimes X) D_m^+ \frac{\partial}{\partial v(X)}) \right] \\ &= \Delta_X^*\end{aligned}$$

Thus, $\Delta_Z^* = \Delta_X^*$ and therefore Δ_X^* is invariant under congruent transformations.

Now, What is the relationship between Δ_Y^* and Δ_X^* ? To see this consider the spectral decomposition of X , i.e. $X = HYH'$ with $H \in O_m$ and $Y = \text{diag}(y_1, \dots, y_m)$. Then

$$\begin{aligned}
(ds)^2 &= \text{tr}(X^{-1}dX X^{-1}dX) \\
&= \text{tr}((HYH')^{-1}d(HYH') (HYH')^{-1}d(HYH')) \\
&= \text{tr}(HY^{-1}H'(dHYH' + HdYH' + HYdH')HY^{-1}H'(dHYH' + HdYH' \\
&\hspace{18em}+HYdH')) \\
&= \text{tr}(H'dHH'dH + dYY^{-1}H'dH + dHH'Y^{-1}H'dHY + Y^{-1}H'dHdY \\
&\hspace{4em}+Y^{-1}dYY^{-1}dY + dHHY^{-1}dY + Y^{-1}H'dHYdH'H \\
&\hspace{4em}+Y^{-1}dYdH'HdY + Y^{-1}H'dH + HdHHdH')
\end{aligned}$$

but $H'dH = -dH'H$ so

$$(ds)^2 = 2\text{tr}(H'dHH'dH) + \text{tr}(Y^{-1}dYY^{-1}dY) - 2\text{tr}Y^{-1}H'dHYH'dH$$

denoting $H'dH = d\Theta$ we have

$$(ds)^2 = 2\text{tr}(d\Theta d\Theta) + \text{tr}(Y^{-1}dYY^{-1}dY) - 2\text{tr}Y^{-1}d\Theta Y d\Theta$$

vectorizing and noting that for a skew-symmetric matrix A , $\tilde{D}_m \tilde{\mathbf{v}}(A) = \text{vec } A$ (equation (6.8), p. 94, Magnus (12)) and for D diagonal $\psi'_m \mathbf{w}(D) = \text{vec}(D)$, (see p. 109 in Magnus (12)).

$$\begin{aligned}
(ds)^2 &= d\text{vec}' Y(Y^{-1} \otimes Y^{-1})d\text{vec } Y - 2d\text{vec}' \Theta(Y \otimes Y^{-1})d\text{vec } \Theta \\
&\hspace{18em}+ 2d\text{vec}' \Theta(I \otimes I)d\text{vec } \Theta \\
&= d\mathbf{w}'(Y)\psi_m(Y^{-1} \otimes Y^{-1})\psi'_m d\mathbf{w}(Y) - 2d\tilde{\mathbf{v}}'(\Theta)\tilde{D}'_m(Y \otimes Y^{-1})\tilde{D}_m \mathbf{v}(\Theta) \\
&\hspace{18em}+ 2d\tilde{\mathbf{v}}'(\Theta)\tilde{D}'_m \tilde{D}_m d\tilde{\mathbf{v}}(\Theta) \\
&= d\mathbf{w}'(Y)\psi_m(Y^{-1} \otimes Y^{-1})\psi'_m d\mathbf{w}(Y) - 2d\tilde{\mathbf{v}}'(\Theta)\tilde{D}'_m((Y \otimes Y^{-1}) - I_{m^2})\tilde{D}_m d\tilde{\mathbf{v}}(\Theta) \\
&= (d\mathbf{w}'(Y) \ d\tilde{\mathbf{v}}'(\Theta)) \begin{pmatrix} \psi_m(Y^{-1} \otimes Y^{-1})\psi'_m & 0 \\ 0 & -2\tilde{D}'_m((Y \otimes Y^{-1}) - I_{m^2})\tilde{D}_m \end{pmatrix} \begin{pmatrix} d\mathbf{w}(Y) \\ d\tilde{\mathbf{v}}(\Theta) \end{pmatrix} \\
&= (d\mathbf{w}'(Y) \ d\tilde{\mathbf{v}}'(\Theta))G(\mathbf{w}(Y)) \begin{pmatrix} d\mathbf{w}(Y) \\ d\tilde{\mathbf{v}}(\Theta) \end{pmatrix}.
\end{aligned}$$

Thus,

$$\det G(\mathbf{w}(Y)) = \det(\psi_m(Y^{-1} \otimes Y^{-1})\psi'_m) \det(-2\tilde{D}'_m((Y \otimes Y^{-1}) - I_{m^2})\tilde{D}_m)$$

note that (see Theorem 7.7(ii), p. 113 in Magnus (12)),

$$\psi_m(Y^{-1} \otimes Y^{-1})\psi'_m = Y^{-1} \odot Y^{-1} = \begin{cases} y_i^{-2} & i = j \\ 0 & i \neq j, \end{cases}$$

where \odot denotes the Hadamard product. Then

$$\det(\psi_m(Y^{-1} \otimes Y^{-1})\psi'_m) = \prod_{i=1}^m y_i^{-2},$$

and from Theorem 6.2(iii), p. 95 in Magnus (12), $\tilde{D}_m^+ = 2\tilde{D}'_m$, then

$$\begin{aligned} \det(-2\tilde{D}'_m((Y \otimes Y^{-1}) + I_{m^2})\tilde{D}_m) &= \det\left(-4\tilde{D}_m^+((Y \otimes Y^{-1}) - I_{m^2})\tilde{D}_m\right) \\ &= -4^{m(m-1)/2} \det\left(\tilde{D}_m^+((Y \otimes Y^{-1}) - I_{m^2})\tilde{D}_m\right) \\ &= -4^{m(m-1)/2} \det\left(\tilde{D}_m^+(Y \otimes Y^{-1})\tilde{D}_m - \tilde{D}_m^+\tilde{D}_m\right), \end{aligned}$$

because $\det(aA) = a^m \det(A)$ if $A \in \mathbb{R}^{m \times m}$. By Theorem 6.15, p. 103 in Magnus (12),

$$\tilde{D}_m^+(Y \otimes Y^{-1})\tilde{D}_m = \text{diag}(1/2(y_i y_j^{-1} + y_j y_i^{-1}), 1 \leq j < i \leq m)$$

and for Theorem 6.2(ii) and (iii), p. 95, from the above reference:

$$\tilde{D}_m^+\tilde{D}_m = 1/2\tilde{D}'_m\tilde{D}_m = 1/2(2I_{m(m-1)/2}) = I_{m(m-1)/2}.$$

Thus

$$\begin{aligned} \det\left(-2\tilde{D}'_m((Y \otimes Y^{-1}) - I)\tilde{D}_m\right) &= -4^{m(m-1)/2} \det\left(\text{diag}\left(\frac{y_i y_j^{-1} + y_j y_i^{-1} - 2}{2}\right), 1 \leq j < i \leq m\right) \\ &= -2^{m(m-1)/2} \det\left(\text{diag}\left(y_i y_j^{-1} + y_j y_i^{-1} - 2\right), 1 \leq j < i \leq m\right) \\ &= -2^{m(m-1)/2} \det\left(\text{diag}\left(\frac{y_i^2 + y_j^2}{y_i y_j} - 2\right), 1 \leq j < i \leq m\right) \\ &= -2^{m(m-1)/2} \det\left(\text{diag}\left(\frac{(y_i^2 - y_j^2)^2}{y_i y_j}\right), 1 \leq j < i \leq m\right) \\ &= \prod_{i < j} 2 \frac{(y_j^2 - y_i^2)^2}{y_i y_j} \end{aligned}$$

then $G(\mathbf{w}(Y))$ is given by

$$G(\mathbf{w}(Y)) = \begin{pmatrix} y_1^{-2} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & y_m^{-2} & 0 & \dots & 0 \\ 0 & \dots & 0 & 2 \frac{(y_2^2 - y_1^2)^2}{y_1 y_2} & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 2 \frac{(y_{m-1}^2 - y_m^2)^2}{y_{m-1} y_m} \end{pmatrix}.$$

Or directly

$$\begin{aligned}\psi_m(Y^{-1} \otimes Y^{-1})\psi'_m &= Y^{-1} \odot Y^{-1} \\ &= \text{diag}(y_1^{-2}, \dots, y_m^{-2})\end{aligned}$$

and

$$\begin{aligned}-2\tilde{D}'_m((Y \otimes Y^{-1}) - I)\tilde{D}_m &= -2\tilde{D}'_m(Y \otimes Y^{-1})\tilde{D}_m + 2\tilde{D}'_m\tilde{D}_m \\ &= -4\tilde{D}_m^+(Y \otimes Y^{-1})\tilde{D}_m + 4I_{m(m-1)/2} \\ &= -4(\text{diag}(1/2(y_i y_j^{-1} + y_j y_i^{-1}), 1 \leq j < i \leq m) - 1) \\ &= -2 \text{diag} \left(\left(\frac{(y_i^2 - y_j^2)^2}{y_i y_j} \right), 1 \leq j < i \leq m \right) \\ &= \text{diag} \left(\left(\frac{2(y_j^2 - y_i^2)^2}{y_i y_j} \right), 1 \leq j < i \leq m \right)\end{aligned}$$

Thus

$$G^{-1}(w(Y)) = \text{diag} \left(y_1^2 \dots y_m^2 \frac{y_1 y_2}{2(y_2^2 - y_1^2)^2} \dots \frac{y_{m-1} y_m}{2(y_{m-1}^2 - y_m^2)^2} \right)$$

then equation (32) of Muirhead (14) can be obtained after the substitution in the operator, however the coefficient of the last term in Muirhead's equation (32) must be changed from 1/4 to 1/2; note that this error does not affect the application of the operator to the zonal polynomials and the last equality in equation (35) of Muirhead (14) holds and it is in agreement with equation (4.5) of James (11). With the recurrence relations, the coefficients of zonal polynomials of positive definite matrix argument can be obtained straightforwardly as Muirhead (14) and James (11) proceed.

4. Zonal Polynomials of positive semidefinite matrix argument: using the Moore-Penrose Inverse. According to the definition of zonal polynomials of symmetric matrix argument, these polynomials are eigenfunctions of the Laplace-Beltrami operator Δ .

Now suppose that $X \in S_m^+(q)$, is an $m \times m$ matrix semidefinite positive of rank q . The following properties hold for the Moore-Penrose inverse X^+ , XX^+ and X^+X are symmetric, $XX^+X = X$ and $X^+XX^+ = X^+$.

Define

$$(ds)^2 = \text{tr}(X^+ dX X^+ dX)$$

If L is orthogonal, ($L \in \mathcal{O}(m)$) under the transformation $X \rightarrow LXL'$ we have

$$\begin{aligned}(ds)^2 &= \text{tr}((LXL')^+ d(LXL')(LXL')^+ d(LXL')) \\ &= \text{tr}(LX^+L' L dXL' LX^+L' L dXL') \\ &= \text{tr}(X^+ dX X^+ dX)\end{aligned}$$

because if $A, C \in \mathcal{O}(m)$ and B is an arbitrary $m \times m$ matrix, $(ABC)^+ = C' B^+ A'$, and $A^+ = A^{-1} = A'$.

Now, it is not possible to use linear structures because if $X \in S_m^+(q)$, is such that

$$X = \begin{pmatrix} X_{11} & X_{12} \\ q \times q & q \times m-q \\ X_{21} & X_{22} \\ m-q \times q & m-q \times m-q \end{pmatrix}, \text{ and } X_{22} = X_{21}X_{11}^{-1}X_{12},$$

then in X only exist $mq - q(q-1)/2$ elements mathematically independent, conformed by the elements mathematically independent in X_{11} , (noting that $X_{11} = X'_{11}$) and the elements in X_{12} . Denote by $\mathbf{u}(X)$ the vector

$$x_{11}, x_{12}, x_{22}, x_{13}, x_{23}, x_{33}, \dots, x_{1q}, x_{2q}, \dots, x_{qq}, x_{1 \ q+1}, x_{2 \ q+1}, \dots, x_{q \ q+1} \\ \dots, x_{1m}, x_{2m}, \dots, x_{qm} \in \mathbb{R}^{mq-q(q-1)/2},$$

we have that

$$\begin{aligned} (ds)^2 &= \text{tr}(X^+ dX X^+ dX) \\ &= d \text{vec}' X (X^+ \otimes X^+) d \text{vec} X \\ &= d \mathbf{u}'(X) \nabla'_m (X^+ \otimes X^+) \nabla_m d \mathbf{u}(X) \end{aligned}$$

with $\nabla_m \in \mathbb{R}^{m^2 \times mq-q(q-1)/2}$ such that $\nabla_m d \mathbf{u}(X) = \text{vec} X$ and

$$G(\mathbf{u}(X)) = \nabla'_m (X^+ \otimes X^+) \nabla_m$$

Unfortunately it does not exist an explicit known form for ∇_m , moreover ∇_m is non unique.

Consider the spectral decomposition of X , i.e. $X = H_1 Y H'_1$ where $H_1 \in V_{q,m}$, $Y = \text{diag}(y_1, \dots, y_q)$, noting that $X^+ = H_1 Y^{-1} H'_1$, then

$$\begin{aligned} (ds)^2 &= \text{tr}(X^+ dX X^+ dX) \\ &= \text{tr}(H_1 Y^{-1} H'_1 (dH_1 Y H'_1 + H_1 dY H'_1 + H_1 Y dH'_1) \\ &\quad H_1 Y^{-1} H'_1 (dH_1 Y H'_1 + H_1 dY H'_1 + H_1 Y dH'_1)) \\ &= \text{tr}(H'_1 dH_1 H'_1 dH_1 + dY Y^{-1} H'_1 dH_1 + dH_1 H'_1 Y^{-1} H'_1 dH_1 Y \\ &\quad + Y^{-1} H'_1 dH_1 dY + Y^{-1} dY Y^{-1} dY \\ &\quad + dH_1 H_1 Y^{-1} dY + Y^{-1} H'_1 dH_1 Y dH'_1 H_1 + Y^{-1} dY dH'_1 H_1 dY + Y^{-1} H'_1 dH_1 \\ &\quad + H_1 dH_1 H_1 dH'_1) \end{aligned}$$

but $H'_1 dH_1 = -dH'_1 H_1 = -(H'_1 dH_1)'$ so

$$(ds)^2 = 2 \text{tr}(H'_1 dH_1 H'_1 dH_1) + \text{tr}(Y^{-1} dY Y^{-1} dY) - 2 \text{tr} Y^{-1} H'_1 dH_1 Y H'_1 dH_1$$

denoting $H'_1 dH_1 = d\Theta_1 \in \mathbb{R}^{q \times q}$ skew-symmetric, we get

$$(ds)^2 = \text{tr}(Y^{-1} dY Y^{-1} dY) - 2 \text{tr}(Y^{-1} d\Theta_1 Y d\Theta_1) + 2 \text{tr}(d\Theta_1 d\Theta_1)$$

vectorizing, given Y diagonal and Θ_1 skew symmetric

$$\begin{aligned}
(ds)^2 &= d \operatorname{vec}' Y (Y^{-1} \otimes Y^{-1}) d \operatorname{vec} Y - 2 d \operatorname{vec}' \Theta_1 (Y \otimes Y^{-1}) d \operatorname{vec} \Theta_1 + 2 d \operatorname{vec}' \Theta_1 \\
&= d \mathbf{w}'(Y) \psi_q (Y^{-1} \otimes Y^{-1}) \psi_q' d \mathbf{w}(Y) - 2 d \tilde{\mathbf{v}}'(\Theta_1) \tilde{D}'_q (Y \otimes Y^{-1}) \tilde{D}_q \mathbf{v}(\Theta_1) \\
&\quad + 2 d \tilde{\mathbf{v}}'(\Theta_1) \tilde{D}'_q \tilde{D}_q d \tilde{\mathbf{v}}(\Theta_1) \\
&= (d \mathbf{w}'(Y), d \tilde{\mathbf{v}}'(\Theta_1)) \begin{pmatrix} \psi_q (Y^{-1} \otimes Y^{-1}) \psi_q' & 0 \\ 0 & -2 \tilde{D}'_q (Y \otimes Y^{-1}) \tilde{D}_q + I_{q^2} \end{pmatrix} \\
&\quad \begin{pmatrix} d \mathbf{w}(Y) \\ d \tilde{\mathbf{v}}(\Theta_1) \end{pmatrix}
\end{aligned}$$

Note that this is the same expression derived in the positive definite case just changing m by q . Then

$$G(\mathbf{w}(Y)) = \operatorname{diag} \left(y_1^{-2} \dots y_q^{-2} \frac{2(y_2^2 - y_1^2)^2}{y_1 y_2} \dots \frac{2(y_{q-1}^2 - y_q^2)^2}{y_{q-1} y_q} \right)$$

Thus $\Delta_X^* = \Delta_{H_1 Y H_1'}^*$ can be written as (noting that the coefficient of the last term is $1/2$ instead of $1/4$ in the analogous equation (32) of Muirhead (14))

$$\begin{aligned}
\Delta_X^* = \Delta_{H_1 Y H_1'}^* &= \sum_{i=1}^q y_i^2 \frac{\partial^2}{\partial y_i^2} + \sum_{i=1}^q \sum_{\substack{j=1 \\ (j \neq i)}}^q \frac{y_i^2}{(y_i - y_j)} \frac{\partial}{\partial y_i} - \frac{1}{2} (q-3) \sum_{i=1}^q y_i \frac{\partial}{\partial y_i} \\
&\quad + \frac{1}{2} \sum_{i < j}^q \frac{y_i y_j}{(y_i - y_j)^2} \frac{\partial^2}{\partial \theta_{ij}^2}, \tag{1}
\end{aligned}$$

with $\Theta_1 = (\theta_{ij})$. Or in terms of the Laplace-Beltrami operator Δ_Y we get

$$\Delta_X^* = \Delta_{H_1 Y H_1'}^* = \Delta_Y - \frac{1}{2} (q-3) \sum_{i=1}^q y_i \frac{\partial}{\partial y_i} + \frac{1}{2} \sum_{i < j}^q \frac{y_i y_j}{(y_i - y_j)^2} \frac{\partial^2}{\partial \theta_{ij}^2}. \tag{2}$$

Now, from the definition of zonal polynomials, the polynomials $C_\kappa(Y)$ are symmetric and homogeneous in the latent roots of Y , and can be expressed as linear combinations of some basic set of symmetric functions in the y_i 's. The following method computes the coefficients required to write the zonal polynomials as linear combinations of the monomial symmetric functions,

$$M_\lambda = y_1^{l_1} y_2^{l_2} \dots y_q^{l_q} + \text{symmetric terms} \tag{3}$$

(namely, over all different permutations of the indices, $\lambda = (l_1, l_2, \dots, l_q)$ in non-increasing order). Thus, the procedure calculates the coefficients c_λ in ,

$$Z_\kappa(Y) = \sum_\lambda c_\lambda M_\lambda \tag{4}$$

where the zonal polynomials are denoted by $Z_\kappa(Y)$ because they are given a different normalizing constant and the partition λ runs through all non-increasing partitions of k into q or fewer parts, and of no higher order than κ . By higher order we are referring to the lexicographical ordering of partitions whereby $\kappa = (k_1, k_2, \dots, k_q)$ is said to be of higher order than $\lambda = (l_1, l_2, \dots, l_q)$ if $k_i > l_i$ for some i , when $k_j = l_j$ for $j = 1, 2, \dots, i - 1$. The zonal polynomials $Z_\kappa(Y)$'s are related to the zonal polynomials $C_\kappa(Y)$, which we have used here, by equations (18) and (19) of James (10):

$$C_\kappa(Y) = c_\kappa y_1^{k_1} y_2^{k_2} \dots y_p^{k_p} + \text{terms of lower weight} \quad (5)$$

$$= \left[\frac{\prod_{i < j}^q (2k_i - 2k_j - i + j)}{2^k k! \prod_{i=1}^q (2k_i + q - i)!} \right] Z_\kappa(Y) \quad (6)$$

where $\kappa = (k_1 \ k_2 \ \dots \ k_q)$ is an ordered partition of k .

On the other hand, given that the real zonal polynomials are functions only of the latent roots, we have that $C_\kappa(Y) = C_\kappa(X)$. Applying the *Euler Operator*, $\sum_{i=1}^q y_i \frac{\partial}{\partial y_i}$, used in (1), to $C_\kappa(Y)$ having the form given in (5), we get

$$\sum_{i=1}^q y_i \frac{\partial}{\partial y_i} C_\kappa(Y) = k C_\kappa(Y), \quad (7)$$

meaning the zonal polynomials are eigenfunctions of the Euler operator with eigenvalue k , in fact any homogeneous polynomial of degree k is an eigenfunction of that operator, with eigenvalue k .

Similarly, if we apply the Laplace-Beltrami operator Δ_Y , from (1) to (7), to the zonal polynomial $C_\kappa(Y)$ given in (5), we have

$$\Delta_Y C_\kappa(Y) = \left[\sum_{i=1}^q k_i (k_i + q - i - 1) \right] C_\kappa(Y), \quad (8)$$

Note that, when the operator $\Delta_X^* = \Delta_{H_1 Y H_1}^*$ in (1) is applied to the zonal polynomial $C_\kappa(Y)$ we obtain the respective eigenvalue ς to be

$$\varsigma = \sum_{i=1}^q k_i (k_i - i) + \frac{1}{2} k (q + 1),$$

where we have used (8) and (5).

Let us call

$$\rho_\kappa = \sum_{i=1}^q k_i (k_i - i).$$

Thus the zonal polynomials of positive semidefinite matrix argument satisfies the partial differential equation

$$\sum_{i=1}^q y_i^2 \frac{\partial^2}{\partial y_i^2} C_\kappa(Y) + \sum_{i=1}^q \sum_{\substack{j=1 \\ (j \neq i)}}^q y_i^2 (y_i - y_j)^{-1} \frac{\partial}{\partial y_i} C_\kappa(Y) = [\rho_\kappa + k(q-1)] C_\kappa(Y) \quad (9)$$

Remark 1. One question arises: Would it be possible to use another kind of generalised inverse instead of the Moore-Penrose inverse, in such a way that the metric form $(ds)^2$ be unique? The answer will be given in Section 7.

5. The Recurrence Relations for the Zonal Polynomials. Now, if κ is a partition of k , then by (3), (4) and (5) the zonal polynomials defined in terms of monomial symmetric functions M_λ are given by

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

where $c_{\kappa, \lambda}$ are constants and the summation is over all partitions λ of k and $\lambda \leq \kappa$, in the sense of lexicographical order explained in last section (see James (10)).

Then the differential equation (9) for zonal polynomials of positive semidefinite matrix argument, in terms of ρ and M_λ is expressed as

$$\begin{aligned} \sum_{i=1}^q y_i^2 \frac{\partial^2}{\partial y_i^2} \sum_{\lambda \leq \kappa} c_{\kappa, \lambda} M_\lambda(Y) + \sum_{i=1}^q \sum_{\substack{j=1 \\ (j \neq i)}}^q y_i^2 (y_i - y_j)^{-1} \frac{\partial}{\partial y_i} \sum_{\lambda \leq \kappa} c_{\kappa, \lambda} M_\lambda(Y) \\ = [\rho_\kappa + k(q-1)] \sum_{\lambda \leq \kappa} c_{\kappa, \lambda} M_\lambda(Y). \end{aligned}$$

Exactly as in James (10), we can derive recurrence relationships for the $c_{\kappa, \lambda}$'s, resulting in

$$c_{\kappa, \lambda} = \frac{1}{\rho_\kappa - \rho_\lambda} \sum_{\lambda < \mu \leq \kappa} [(l_i + r) - (l_j - r)] c_{\kappa, \mu},$$

and the notations holds for

$$\begin{aligned} \lambda &= (l_1, \dots, l_q) \\ \mu &= (l_1, \dots, l_i + r, \dots, l_j - r, \dots, l_q) \end{aligned}$$

for all r such that at first μ is not necessarily a lexicographic order partition but its final form must be in descending order and $\lambda < \mu \leq \kappa$.

6. Calculations of the Zonal Polynomials. Knowing that zonal polynomials of positive definite and semidefinite matrix argument, have the same recurrence relation when the Moore-Penrose inverse is used in the last case, will permit us to do computations in exactly the same way as in James (1968). In this way, we can use the available tables

for zonal polynomials up to 12th degree (see Parkhurst and James (15)) just changing the complete rank m for the rank $q < m$ of X . Besides, a computational algorithm given by McLaren (13) can be applied here to calculate the zonal polynomials in the semidefinite case as it does in the definite case.

7. Other Inverses. We now consider the question posed at the end of Section 4, but first we give a brief summary of some generalised inverses. If $A = A' \geq 0$, is an $m \times m$ matrix, then its spectral decomposition is

$$A = H \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} H'$$

where $H \in \mathcal{O}(m)$, $\Delta = \text{diag}(\delta_1, \dots, \delta_r)$ with $r(A) = r(\Delta) = r$, then

$$A^+ = H \begin{pmatrix} \Delta^{-1} & 0 \\ 0 & 0 \end{pmatrix} H', \quad \text{Moore-Penrose inverse} \quad (1)$$

$$A^- = H \begin{pmatrix} \Delta^{-1} & N \\ N' & M \end{pmatrix} H', \quad \text{Generalised inverse or } g\text{-inverse} \quad (2)$$

$$A^r = H \begin{pmatrix} \Delta^{-1} & N \\ N' & N'MN \end{pmatrix} H', \quad \text{Reflexive } g\text{-inverse} \quad (3)$$

$$A^l = H \begin{pmatrix} \Delta^{-1} & 0 \\ 0 & M \end{pmatrix} H', \quad \text{Least squares } g\text{-inverse} \quad (4)$$

$$A^m = A^l, \quad \text{Minimum norm } g\text{-inverse} \quad (5)$$

where N and M arbitrary matrices of proper orders, see Rao (16), p.76-77, problem 28. Now, let A^g denote any generalised inverse, and consider the properties

$$AA^g \text{ is symmetric} \quad (6)$$

$$A^g A \text{ is symmetric} \quad (7)$$

$$AA^g A = A \quad (8)$$

$$A^g AA^g = A^g \quad (9)$$

Then, A^+ satisfies 6 to 9; A^- satisfies 8; A^r satisfies 8 and 9; A^l satisfies 6 and 8 and A^m satisfies 7 and 8.

Let us see the behavior of $(ds)^2$ under A^l , A^m , A^- and A^r . With this intention, denote generically A^l , A^m , A^- and A^r by

$$A^g = H \begin{pmatrix} \Delta^{-1} & N \\ N' & K \end{pmatrix} H' = H\Gamma H' \quad (10)$$

where $K = M$ or $K = N'MN$ according to the generalised inverse A^- or A^r taken or $N = 0$ and $K = M$ in the cases of A^l and A^m . Also, denote

$$A = H \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} H' = HRH' \quad (11)$$

Then the metric has the form

$$(ds)^2 = \text{tr}(X^g dX X^g dX) \quad (12)$$

First, observe that $(ds)^2$, defined in (12), is invariant under the congruence transformation $X \rightarrow LXL'$ ($L \in Gl(m, \mathbb{R})$) for the A^- and A^r cases. And $(ds)^2$ is invariant under the transformation $X \rightarrow LXL'$, $L \in \mathcal{O}(m)$ for the A^l and A^m cases.

Now, considering (11) and (10), we have

$$\begin{aligned} (ds)^2 &= \text{tr}(X^g dX X^g dX) \\ &= \text{tr}(H\Gamma H' (dHRH' + HdRH' + HRdH') \\ &\quad H\Gamma H' (dHRH' + HdRH' + HRdH')) \\ &= \text{tr}((H\Gamma H' dHRH' H\Gamma H' + H\Gamma H' HdRH' H\Gamma H' + H\Gamma H' HRdH' H\Gamma H') \\ &\quad (dHRH' + HdRH' + HRdH')) \end{aligned} \quad (13)$$

Working as in Section 4 and considering that

$$B = \Gamma R = \begin{pmatrix} \Delta' & N \\ N' & K \end{pmatrix} \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ N'\Delta & 0 \end{pmatrix} \quad (14)$$

And,

$$R\Gamma = \begin{pmatrix} \Delta & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta' & N \\ N' & K \end{pmatrix} = \begin{pmatrix} I & \Delta N \\ 0 & 0 \end{pmatrix} = (\Gamma R)' \quad (15)$$

We have that

$$(ds)^2 = 2 \text{tr}(B' H' dHB' H' dH) + \text{tr}(\Gamma dR\Gamma dR) - 2 \text{tr}(\Gamma H' dHRH' dH) \quad (16)$$

Now using the fact that $B'R = R = RB$ and writing $H'dH = d\Theta$, we have that

$$(ds)^2 = 2 \text{tr}(B' d\Theta B' d\Theta) + \text{tr}(\Gamma dR\Gamma dR) - 2 \text{tr}(\Gamma d\Theta R d\Theta) \quad (17)$$

Note that $(ds)^2$ defined in (12) depends of B and R , which depend of the arbitrary matrices N and K , thus $(ds)^2$ defined in this way is not unique. Thus the only way in constructing the zonal polynomials of positive semidefinite matrix argument is by the Moore-Penrose inverse, in which case it was proved, the metric $(ds)^2$ is unique.

8. Acknowledgment. This work was supported partially by the research project 39017E of CONACYT-México. Also, thanks to Dr. Rogelio Ramos Quiroga (CIMAT-México) for his careful reading and excellent comments.

REFERENCES

- [1] CONSTANTINE, A. G. (1963). Noncentral distribution problems in multivariate analysis. *Annals of Mathematical Statistics*, **34**, 1270–1285.
- [2] CONSTANTINE, A. G. (1966). The distribution of Hotelling's generalized T_0^2 . *Annals of Mathematical Statistics*, **37**, 215–225.

- [3] DÍAZ-GARCÍA, J. A., GUTIÉRREZ J. R., AND MARDIA, K. V. (1997). Wishart and Pseudo-Wishart distributions and some applications to shape theory. *Journal of Multivariate Analysis* **63**, 73-87.
- [4] DÍAZ-GARCÍA, J. A., AND GONZÁLEZ-FARÍAS, G. (2004). Singular Random Matrix Decompositions: Distributions. *Journal of Multivariate Analysis* to appear.
- [5] FARRELL, R. H. (1985). *Multivariate Calculation: Use of the Continuous Groups*. Springer Series in Statistics, Springer-Verlag, New York.
- [6] HELGASON, S. (1962). *Differential Geometry and Symmetric Spaces*. Academic Press, New York.
- [7] JAMES, A. T. (1960). The distribution of the latent roots of the covariance matrix. *Annals of Mathematical Statistics*, **31**, 151–158.
- [8] JAMES, A. T. (1961a). Zonal polynomials of the real positive definite symmetric matrices. *Annals of Mathematics*, **74**(2), 456–469.
- [9] JAMES, A. T. (1961b). The distribution of noncentral means with known covariance. *Annals of Mathematical Statistics*, **32**, 874–882.
- [10] JAMES, A. T. (1964). Distributions of matrix variate and latent roots derived from normal samples. *Annals of Mathematical Statistics*, **35**, 475–501.
- [11] JAMES, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator. *Annals of Mathematical Statistics*, **39**, 1711–1718.
- [12] MAGNUS, J. R. (1988). *Linear Structures.*, Charles Griffin & Company Ltd, London.
- [13] MCLAREN, M. (1976). Algorithm AS 94: Coefficients of the Zonal Polynomials. *Applied Statistics*, **25-1**, 82-87.
- [14] MUIRHEAD, R. J. (1982). *Aspects of multivariate statistical theory*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.
- [15] PARKHURST, A. M., AND JAMES, A. T. (1974). Zonal polynomials of order 1 to 12. *Selected tables in mathematical statistics, Vol. 2*, American Mathematical Society, Providence.
- [16] RAO, C. R. (1973). *Linear Statistical Inference and its Applications*. Second Edition, John Wiley & Sons, Inc., New York.
- [17] TAKEMURA, A. (1984). *Zonal polynomials*. Lecture Notes-Monograph Series, **4**, Institute of Mathematical Statistics, Hayward, California.

JOSÉ A. DÍAZ-GARCÍA
DEPARTMENT OF STATISTICS AND
COMPUTATION
UNIVERSIDAD AUTÓNOMA AGRARIA
ANTONIO NARRO
25350 BUENAVISTA, SALTILLO,
COAHUILA, MÉXICO
E-MAIL: jadiaz@uaaan.mx

FRANCISCO JOSÉ CARO
DEPARTMENT OF MATHEMATICS
UNIVERSIDAD DE ANTIOQUIA
MEDELLÍN, A. A. 1226
COLOMBIA
E-MAIL: fjcaro@matematicas.udea.edu.co