3 C. R. Johnson, Some outstanding problems in the theory of matrices, *Linear and Multilinear Algebra* 12:99–108 (1982).
4 T. Laffey, Simultaneous reduction of sets of matrices under similarity, *Linear Algebra Appl.* 84:123–138 (1986).
5 Z. P. Lei, Permutation equivalence and permutation similarity of matrices, *Math. Practice Theory* 1:34–43 (1983).
6 J. F. Maitre and N. H. Vinh, Valeurs singulières généralisées et meilleure approximation de rang r d'un opérateur linéaire, *C.R. Acad. Sci. Paris* 262-A:502–504 (1966).
7 H. Schneider and R. Turner, Matrices Hermitian for an absolute norm, *Linear and Multilinear Algebra*, 1:9–31 (1973).
8 M. J. Sodupe, Thesis Doctoral, Bilbao, 1987.
9 I. Vidav, The group of isometries and the structure of a finite-dimensional Banach space, *Linear Algebra Appl.* 14:227–236 (1976).

AN APPLICATION OF ZONAL POLYNOMIALS TO THE GENERATION OF PROBABILITY DISTRIBUTIONS

by R. GUTIÉRREZ JÁIMEZ and J. A. HERMOSO GUTIÉRREZ[17]

1. *Introduction*

The group representation theory of $\mathrm{Gl}(m, R)$, the general linear group, can be used to define the zonal polynomials. From a technical point of view it is quite difficult to define the zonal polynomials. The group-theoretic construction of zonal polynomials can be seen in Farrell (1976) and James (1961, 1964, 1968, 1973, 1976).

The approach to zonal polynomials through group representation theory is as follows. Let V_k be the vector space of homogeneous polynomials $\Phi(X)$ of degree k in the $n = m(m+1)/2$ different elements of the $m \times m$ symmetric matrix X. Corresponding to any congruence transformation $X \to LXL'$, $L \in \mathrm{Gl}(m, R)$, a linear transformation of the space V_k can be defined by $\Phi \to T(L)\Phi : (T(L)\Phi)(X) = \Phi(L^{-1}XL^{-1\prime})$. A representation of $\mathrm{Gl}(m, R)$ in the vector space V_k is defined by this transformation, that is, the mapping $L \to T(L)$ is a homomorphism from $\mathrm{Gl}(m, R)$ to the group of linear transformations of V_k, $T(L_1L_2) = T(L_1)T(L_2)$.

[17]Departamento de Estadística, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain.

A subspace $V' \subset V_k$ is said to be invariant if $T(L)V' \subset V'$ for all $L \in \mathrm{Gl}(m, R)$. If V' does not contain proper invariant subspaces, it is called an irreducible invariant subspace. This is the way in which the zonal polynomials arise. It can be proved that the space V_k decomposes into a direct sum of irreducible invariant subspaces V_κ, $V_k = \oplus_\kappa V_\kappa$, where $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_m)$, $\kappa_1 \geqslant \kappa_2 \geqslant \cdots \geqslant \kappa_m \geqslant 0$, runs over all partitions of k into not more than m parts.

The polynomial $(\operatorname{tr} X)^k \in V_k$ has a unique decomposition $(\operatorname{tr} X)^k = \Sigma_\kappa C_\kappa(X)$ into polynomials $C_\kappa(X) \in V_\kappa$ belonging to the respective invariant subspaces. The polynomial $C_\kappa(X)$ is the zonal polynomial corresponding to the partition κ.

We will start from another definition for the zonal polynomials. This new definition gives a clear sense to zonal polynomials: they are a generalization of the powers y^k when y is replaced by a matrix Y.

Let Y be an $m \times m$ symmetric matrix with latent roots $y_1, \ldots, y_m$, and let $\kappa = (\kappa_1, \ldots, \kappa_m)$ be a partition of k into not more than m parts. The zonal polynomial 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:

(i) The term of highest weight (lexicographically ordered) in $C_\kappa(Y)$ is $y_1^{\kappa_1} \cdots y_m^{\kappa_m}$, that is, $C_\kappa(Y) = d_\kappa y_1^{\kappa_1} \cdots y_m^{\kappa_m} +$ terms of lower weight.

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

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

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

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

At this point it should be stated that no general formula for zonal polynomials is known; however, the above description provides a general algorithm for their calculation (Muirhead, 1982).

2. *Probability Generating Functions of Matrix Argument*

Power series are often used in statistics, for example the probability generating functions (p.g.f.'s). The well-known hypergeometric probability distribution is generated, except for a constant factor, by the classical (or Gaussian) hypergeometric series

$$ {}_2F_1(a,b;c;y) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k} \frac{y^k}{k!}. \tag{1}$$

The preceding series may be generalized to several variables in different ways, such as the Appell's and Lauricella's functions (Srivastava and Manocha, 1984; Steyn, 1951). By generalizing the powers in (1) by means of zonal polynomials we get probability generating functions of matrix argument,

$$ {}_2F_1(a,b;c;Y) = \sum_{k=0} \sum_{\kappa} \frac{(a)_\kappa(b)_\kappa}{(c)_\kappa} \frac{C_\kappa(Y)}{k!}, \tag{2}$$

where Σ_κ denotes summation over all partitions κ of k, $C_\kappa(Y)$ is the zonal polynomial of Y corresponding to κ, $(a)_\kappa$ is the generalized hypergeometric coefficient, and Y is an $m \times m$ symmetric matrix (Muirhead, 1982, p. 258).

The previous function (2) can be considered, except for a constant factor, as a probability generating function for an m-dimensional discrete random vector $(X_1,\ldots,X_m)$ having a homogeneous probability function, that is, the probabilities $P(X_1=\kappa_1,\ldots, X_m=\kappa_m)$ and $P(X_1=\kappa_{i_1},\ldots, X_m=\kappa_{i_m})$ are all the same, where $(\kappa_{i_1},\ldots,\kappa_{i_m})$ is a permutation of $(\kappa_1,\ldots,\kappa_m)$.

If the general formula for zonal polynomials were known, we would know the probability distribution and, differentiating in the usual way, we could obtain the moments. Because we have no such formula, we must use certain properties of the zonal polynomials and the function (2) to obtain the characteristics of the previously mentioned probability distributions.

Using the integral representation for (2) (Muirhead, 1982),

$$\frac{\Gamma_m(c)}{\Gamma_m(a)\Gamma_m(c-a)} \int_{0<X<I_m} [\det(I-YX)]^{-b} (\det X)^{a-\frac{1}{2}(m+1)}$$

$$\times [\det(I-X)]^{c-a-\frac{1}{2}(m+1)} (dX),$$

we obtain the constant factor:

$$C=\frac{\beta_m(a,c-a)}{\beta_m(a,c-a-b)}=\frac{\Gamma_m(c-a)\Gamma_m(c-b)}{\Gamma_m(c)\Gamma_m(c-a-b)}=\frac{(c-a-b)_{(a,\ldots,a)}}{(c-a)_{(a,\ldots,a)}}. \quad (3)$$

This result is analogous to the one obtained in the univariate case and also for Lauricella's and Appell's functions.

Although the zonal polynomial formula is not generally known, it is known for particular cases, such as $Y=I_m$ (James, 1964). Using (2) and $C_\kappa(I_m)$, the expression for C can be obtained in a different way. That expression and (3) lead to the following relation, which can be used to evaluate (recurrently) the Beta multivariate function:

$$\frac{\beta_m(a,b)}{\beta_m(a,c)}=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a)_\kappa(c-b)_\kappa(\frac{1}{2}m)_\kappa}{(c-a)_\kappa}2^{2k}\frac{\prod_{i<j}^{p}(2\kappa_i-2\kappa_j-i+j)}{\prod_{i=1}^{p}(2\kappa_i+p-i)!}.$$

The function ${}_2F_1(a,b;c;Y)$ satisfies the partial differential equation (p.d.e.)

$$\delta_Y F+\left[c-\tfrac{1}{2}(m-1)\right]\varepsilon_Y F-\Delta_Y F-\left[a+b+1-\tfrac{1}{2}(m-1)\right]E_Y F=mabF$$

expressed in terms of differential operators, where Δ_Y is introduced in the definition of zonal polynomials and

$$E_Y=\sum_{i=1}^{m}y_i\frac{\partial}{\partial y_i},\qquad \varepsilon_Y=\sum_{i=1}^{m}\frac{\partial}{\partial y_i},$$

$$\delta_Y=\sum_{i=1}^{m}y_i\frac{\partial^2}{\partial y_i^2}+\sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\frac{y_i}{y_i-y_j}\frac{\partial}{\partial y_i}.$$

It is possible to prove a much stronger result than the one given above. The function ${}_2F_1(a,b;c;Y)$ is the unique solution of each of the m p.d.e.'s (Muirhead, 1982)

$$y_i(1-y_i)\frac{\partial^2 F}{\partial y_i^2}+\left\{c-\tfrac{1}{2}(m-1)-\left[a+b+1-\tfrac{1}{2}(m-1)\right]y_i\right.$$

$$\left.+\frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_i(1-y_i)}{y_i-y_j}\right\}\frac{\partial F}{\partial y_i}$$

$$-\frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{y_j(1-y_j)}{y_i-y_j}\frac{\partial F}{\partial y_j}=abF,\qquad i=1,\dots,m. \tag{4}$$

Using (4), it can be shown that the moment generating function $M=M(\alpha_1,\dots,\alpha_m)$ is a solution of the following system

$$(1-e^{\alpha_i})\frac{\partial^2 M}{\partial \alpha_i^2}+\left\{c-1-\tfrac{1}{2}(m-1)-\left[a+b-\tfrac{1}{2}(m-1)\right]e^{\alpha_i}\right.$$

$$\left.+\frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{e^{\alpha_i}(1-e^{\alpha_i})}{e^{\alpha_i}-e^{\alpha_j}}\right\}\cdot\frac{\partial M}{\partial \alpha_i}$$

$$-\frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{e^{\alpha_i}(1-e^{\alpha_j})}{e^{\alpha_i}-e^{\alpha_j}}\frac{\partial M}{\partial \alpha_j}=abMe^{\alpha_i},\qquad i=1,\dots,m, \tag{5}$$

where $e^{\alpha_i}=y_i$, $i=1,\dots,m$.

After putting $\alpha_i=0$ in the ith equation of (5), it clearly follows that

$$\left[(c-a-b-1)\frac{\partial M}{\partial \alpha_i}\right]_{\alpha_i=0}=\left[\frac{1}{2}\sum_{\substack{j=1\\j\neq i}}^{m}\frac{\partial M}{\partial \alpha_j}+abM\right]_{\alpha_i=0}.$$

According to a necessary and sufficient condition for rational regression in terms of differential operators (Steyn, 1956), the equation above shows that the regression equation of X_i with respect to the other variables is given by

$$\hat{x}_i = \frac{ab + \frac{1}{2}\sum_{j \neq i} x_j}{c - a - b - 1},$$

and so it is linear (Gutiérrez and Hermoso, 1987).

It follows from (5) that the cumulant generating function, $L = \ln M$, satisfies the system of p.d.e.'s

$$(1 - e^{\alpha_i})\left\{\frac{\partial^2 L}{\partial \alpha_i^2} + \left(\frac{\partial L}{\partial \alpha_i}\right)^2\right\} + \left\{c - 1 - \tfrac{1}{2}(m-1) - \left[a + b - \tfrac{1}{2}(m-1)\right]e^{\alpha_i} + \frac{1}{2}\sum_{\substack{j=1\\ j \neq i}}^{m} \frac{e^{\alpha_i}(1 - e^{\alpha_i})}{e^{\alpha_i} - e^{\alpha_j}}\right\}\frac{\partial L}{\partial \alpha_i}$$

$$- \frac{1}{2}\sum_{\substack{j=1\\ j \neq i}}^{m} \frac{e^{\alpha_i}(1 - e^{\alpha_j})}{e^{\alpha_i} - e^{\alpha_j}}\frac{\partial L}{\partial \alpha_j} = abe^{\alpha_i}, \qquad i = 1, \ldots, m. \tag{6}$$

By successive differentiation of the system (6) we get equations satisfied by the cumulants and then the moments:

$$\bar{x}_i = E[X_i] = \frac{ab}{c - a - b - \frac{1}{2}(m+1)},$$

$$\sigma_i^2 = \frac{(\bar{x}_i + a)(\bar{x}_i + b)}{c - a - b - \frac{1}{2}(m+1) - 1},$$

$$\sigma_{ij} = \frac{\sigma_i^2}{2\left(c - a - b - \frac{1}{2}m\right)}$$

(Gutiérrez and Hermoso, 1987).

These expressions are similar to the expressions of the moments of the distributions generated by Lauricella's function (Steyn, 1951).

The authors thank the referee for helpful comments.

REFERENCES

Constantine, A. G. 1963. Some noncentral distribution problems in multivariate analysis, *Ann. Math. Statist.* 34:1270–1285.

Constantine, A. G. 1966. The distribution of Hotelling's generalized T_0^2, *Ann. Math. Statist.* 37:215–225.

Constantine, A. G. and Muirhead, R. J. 1972. Partial differential equations for hypergeometric functions of two argument matrices, *J. Multivariate Anal.* 3:332–338.

Farrell, R. H. 1976. *Techniques of Multivariate Calculation*. Springer, New York.

Gutiérrez Jáimez, R. and Hermoso Gutiérrez, J. A. 1987. On discrete multivariate probability distributions generated by hypergeometric functions of matrix argument, presented at 17th European Meeting of Statisticians, Thessaloniki, Greece.

James, A. T. 1961. Zonal polynomials of the real positive definite symmetric matrices, *Ann. Math.* 74:456–469.

James, A. T. 1964. Distributions of matrix variates and latent roots derived from normal samples, *Ann. Math. Statist.* 35:475–501.

James, A. T. 1968. Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator, *Ann. Math. Statist.* 39:1711–1718.

James, A. T. 1973. The variance information manifold and the functions on it, in *Multivariate Analysis* (P. R. Krishnaiah, Ed.), Vol. III, Academic, New York, pp. 157–169.

James, A. T. 1976. Special functions of matrix and single argument in statistics, in *Theory and Applications of Special Functions* (R. A. Askey, Ed.), Academic, New York, pp. 497–520.

Muirhead, R. J. 1970. System of partial differential equations for hypergeometric functions of matrix argument, *Ann. Math. Statist.* 41:991–1001.

Muirhead, R. J. 1982. *Aspects of Multivariate Statistical Theory*, Wiley.

Srivastava, H. M. and Manocha, H. L. 1984. *A Treatise on Generating Functions*, Math. Appl., Ellis Horwood.

Steyn, H. S. 1951. On discrete multivariate probability functions, *Kon. Ned. Akad. Wet. Proc. A* 54:23–30.

Steyn, H. S. 1955. On discrete multivariate probability functions of hypergeometric type, *Kon. Ned. Akad. Wet. Proc. A* 58:588–595.

Steyn, H. S. 1956. On regression properties of discrete systems of probability functions, *Kon. Ned. Akad. Wet. Proc. A* 63:119–127.