

***DISTRIBUTION OF THE GENERALISED
INVERSE OF A RANDOM MATRIX AND ITS
APPLICATIONS***

*José A. Díaz-García, Ramón Gutiérrez-Jáimez
and Ramón Gutiérrez-Sánchez*

Comunicación Técnica No I-03-05/10-03-2003
(PE/CIMAT)



Distribution of the generalised inverse of a random matrix and its applications

José A. Díaz-García*

Department of Statistics and Computation
Universidad Autónoma Agraria Antonio Narro
25350 Buenavista, Saltillo, Coahuila, MÉXICO.
jadiaz@uaaan.mx

Ramón Gutiérrez-Jáimez and Ramón Gutiérrez-Sánchez

Department of Statistics and O.R
University of Granada
Granada 18071, Spain.
rgjaimez@goliat.ugr.es and ramongs@goliat.ugr.es

KEY WORD AND PHRASES: Generalised Inverse Wishart and Pseudo-Wishart distributions, Jacobian, Hausdorff measure, More-Penrose inverse, Matrix-variate T singular distribution, Bayesian inference, Matrix-variate Normal singular distribution, Multivariate linear model.

ABSTRACT

Given a random singular matrix X , in the present article we find the Jacobian of the transformation $Y = X^+$, where X^+ is the More-Penrose inverse of X , both in the general case and when X is a non-negative definite matrix. Expressions for the densities of the More-Penrose inverse of the singular Wishart and Pseudo-Wishart matrices are obtained. Similarly, an expression for the density of the matrix-variate singular T -distribution is proposed. Finally, these results are applied to the Bayesian inference of the multivariate linear model.

1. INTRODUCTION

The distribution of the inverse of a random matrix plays an important role in Bayesian inference. Such are the roles of the inverse Wishart, Beta and Dirichlet distributions, see Press [Sections 8.6.1 and 8.6.2, 1982], Box and Tiao [p.460, 1972] and Xu (1990). In particular, the inverse Wishart distribution is obtained as the posterior distribution of the covariance matrix in the inference of the multivariate linear model, when a prior noninformative distribution is assumed.

*This article was written while the first author was a Visiting Professor at the Department of Statistics and O. R. of the University of Granada, Granada, Spain

Similarly, the inverse Wishart distribution is the natural conjugate prior distribution for the covariance matrix in a normal distribution (or a multivariate linear model, with normal errors), in which case this distribution is also obtained as an approximation of the posterior marginal distribution of the matrix of covariances, see Press [pp. 117, 252-253, 256, 1982]. In all these cases, it is assumed that the matrix of covariances is positive definite, although this is not always the case. In order to extend these ideas from the nonsingular to the singular case, it is necessary to find the distributions of the inverse, which in this case would be those of the generalised inverse, or the More-Penrose inverse.

It is necessary to determine the distributions of other singular random matrices. In the above-cited case, firstly, the joint distribution of the errors in the multivariate linear model is a matrix-variate Normal singular distribution; and secondly, by analogy with the nonsingular case, the posterior distribution of the parameters of the linear model would be a matrix-variate singular T -distribution. Several authors have examined these problems, see Uhlig (1994), Díaz-García et al. (1997), Díaz-García and Gutiérrez (1997) and Díaz-García and Gutiérrez-Jáimez (2003).

Note that, given an X singular random matrix, its density function can be written as $dF_X(X) = f_X(X)(dX)$, where (dX) denotes the Hausdorff measure, see Díaz-García et al. (1997) and Billingley [p. 247, 1979]. Let us now assume we wish to find the density of Y , defined by $Y = X^+$, where X^+ is the More-Penrose inverse of X (see Section 2), which would be given by

$$dF_Y(Y) = f_X(Y^+) |J(X \rightarrow Y)| (dY).$$

The problem is reduced to one of calculating the Jacobian $|J(X \rightarrow Y)|$ and of explicitly defining the volume element (dY) . One means that has proved very useful, both for the calculation of Jacobians and for the explicit definition of the Hausdorff measure, is that of the exterior product, see James (1954), Muirhead [Chapter 2, 1982], Uhlig (1994), Díaz-García et al. (1997) and Díaz-García and Gutiérrez (1997).

The present paper extends some results of the above-cited nonsingular case to the singular case. Specifically, Section 2 introduces the notation necessary for the rest of the study and establishes some results with reference to the distribution of singular random matrices. Section 3 presents the Jacobians and the volume elements corresponding to the transformation $Y = X^+$, both for the case of a general singular random matrix and for that of a non-negative defined X random matrix. In Section 4, we determine the singular and nonsingular generalised inverse Wishart and Pseudo-Wishart distributions, both for the central and the noncentral cases. This section also provides an explicit expression for the density of a matrix-variate singular T -distribution. Finally, these results are applied to the Bayesian inference of the multivariate linear model.

2. PRELIMINARY RESULTS AND NOTATION

Let $\mathcal{L}_{m,N}^+(q)$ be the linear space of all $N \times m$ real matrices of rank $q \leq \min(N, m)$ with q distinct singular values. The set of matrices $H_1 \in \mathcal{L}_{m,N}$ such that $H_1' H_1 = I_m$ is a manifold denoted $\mathcal{V}_{m,N}$, called Stiefel manifold. In particular, $\mathcal{V}_{m,m}$ is the group of orthogonal matrices $\mathcal{O}(m)$. Denote \mathcal{S}_m , the homogeneous space of $m \times m$ positive definite symmetric matrices; $\mathcal{S}_m^+(q)$, the $(mq - q(q - 1)/2)$ -dimensional manifold of rank q positive semidefinite $m \times m$ symmetric matrices with q distinct positive eigenvalues.

Observe that, if $X \in \mathcal{L}_{m,N}^+(q)$, we can write X as

$$X_1 = \begin{pmatrix} X_{11} & X_{12} \\ q \times q & q \times m - q \\ X_{21} & X_{11} \\ N - q \times q & N - q \times m - q \end{pmatrix}$$

such that $r(X_{11}) = q$. This is equivalent to the right product of the matrix X with a permutation matrix Π , see Golub and Van Loan [section 3.4.1, 1996], that is $X_1 = X\Pi$. Note that the exterior product of the elements from the differential matrix dX are not affected by the fact that we multiply X (right or left) by a permutation matrix, that is, $(dX_1) = (d(X\Pi)) = (dX)$, since Π is an orthogonal matrix, see Muirhead [Section 2.1, 1982] and James (1954). Then, without loss of generality, (dX) will be defined as the exterior product for the differentials dx_{ij} , such that x_{ij} are mathematically independent. It is important to note that we will have $Nq + mq - q^2$ mathematically independent elements in the matrix $X \in \mathcal{L}_{m,N}^+(q)$, corresponding to the elements of X_{11}, X_{12} and X_{21} . Explicitly,

$$(dX) \equiv (dX_{11}) \wedge (dX_{12}) \wedge (dX_{21}) = \bigwedge_{i=1}^N \bigwedge_{j=1}^q dx_{ij} \bigwedge_{i=1}^q \bigwedge_{j=q+1}^m dx_{ij} \quad (1)$$

Similarly, given $S \in \mathcal{S}_m^+(q)$, we define (dS) as

$$(dS) \equiv \bigwedge_{i=1}^q \bigwedge_{j=i}^m ds_{ij} \quad (2)$$

Again, we should note that, for this case, the matrix S can be written as

$$S \equiv \begin{pmatrix} S_{11} & S_{12} \\ q \times q & q \times m - q \\ S_{21} & S_{22} \\ m - q \times q & m - q \times m - q \end{pmatrix} \quad \text{with} \quad r(S_{11}) = q.$$

such that, the number of mathematically independent elements in S are, $mq - q(q - 1)/2$ corresponding to the mathematically independent elements of S_{12} and S_{11} . Recall that $S_{11} \in \mathcal{S}_q$, in such a way that S_{11} has $q(q + 1)/2$, therefore,

$$(dS) \equiv (dS_{11}) \wedge (dS_{12})$$

Note now that the explicit form for (dX) and (dS) depends on the factorisation (base and coordinate set) employed to represent X or S . By using the nonsingular part of the decomposition in singular values and the nonsingular part of the spectral decomposition for X and S , respectively, then:

Proposition 1. [Singular value decomposition, SVD.] *Let $X \in \mathcal{L}_{m,N}^+(q)$, then there exist $H_1 \in \mathcal{V}_{q,N}$, $P_1 \in \mathcal{V}_{q,m}$ and $D = \text{diag}(D_{11}, \dots, D_{qq})$, $D_{11} > \dots > D_{qq} > 0$, such that $X = H_1 D P_1'$, it is called nonsingular part of the SVD, Rao [p. 42, 1973] and Eaton [p. 58, 1983]. Let $H_2 \in \mathcal{V}_{N-q,N}$ (a function of H_1) and $P_2 \in \mathcal{V}_{m-q,m}$ (a function of P_1) such that $H = (H_1 | H_2) \in$*

$\mathcal{O}(N)$ and $P = (P_1|P_2) \in \mathcal{O}(m)$. Writing by columns, $H_1 = (h_1 \cdots h_q)$, $H_2 = (h_{q+1} \cdots h_N)$, $P_1 = (p_1 \cdots p_q)$ and $P_2 = (p_{q+1} \cdots p_m)$, we have that

$$(dX) = 2^{-q}|D|^{N+m-2q} \prod_{i<j}^q (D_{ii}^2 - D_{jj}^2)(dD)(H_1' dH_1)(P_1' dP_1), \quad (3)$$

where $(dD) \equiv \bigwedge_{i=1}^q dD_{ii}$, and

$$(H_1' dH_1) \equiv \bigwedge_{i=1}^q \bigwedge_{j=i+1}^N h_j' dh_i \quad \text{and} \quad (P_1' dP_1) \equiv \bigwedge_{i=1}^q \bigwedge_{j=i+1}^m p_j' dp_i$$

define an invariant measure on $\mathcal{V}_{q,N}$ and on $\mathcal{V}_{q,m}$, respectively, James (1957) Muirhead [Section 2.1.4] (1982) and Farrell (1985).

For a proof see Díaz-García et al. (1997).

Proposition 2. [Spectral decomposition.] *Let $S \in \mathcal{S}_m^+(q)$, then $S = W_1 L W_1'$, where $W_1 \in \mathcal{V}_{q,m}$ and $L = \text{diag}(L_{11}, \dots, L_{qq})$, $L_{11} > \cdots > L_{qq} > 0$, it is called the nonsingular part of the spectral decomposition. Then*

$$(dS) = 2^{-q}|L|^{m-q} \prod_{i<j}^q (L_{ii} - L_{jj})(dL)(W_1' dW_1) \quad (4)$$

where $(dL) = \bigwedge_{i=1}^q dL_{ii}$, see Uhlig (1994) and Díaz-García and Gutiérrez (1997).

Definition 1. [Matrix-variate Normal Singular Distribution] *Let $X \sim \mathcal{N}_{N \times m}(\mu, \Sigma, \Xi)$, with Σ $m \times m$, $r(\Sigma) = r < m$ or Ξ $N \times N$, $r(\Xi) = k < N$. This distribution will be called a matrix-variate normal singular distribution and will be denoted as*

$$X \sim \mathcal{N}_{N \times m}^{k,r}(\mu, \Sigma, \Xi)$$

omitting the supra-index when $r = m$ and $k = N$. In addition, its density function is given by

$$\frac{1}{(2\pi)^{rk/2} \left(\prod_{i=1}^r \lambda_i^{k/2} \right) \left(\prod_{j=1}^k \delta_j^{r/2} \right)} \text{etr} \left(-\frac{1}{2} \Sigma^- (X - \mu)' \Xi^- (X - \mu) \right) \quad (5)$$

$$\left. \begin{aligned} H_2' X P_1' &= H_2' \mu P_1' \\ H_1' X P_2' &= H_1' \mu P_2' \\ H_2' X P_2' &= H_2' \mu P_2' \end{aligned} \right\} \text{ a. s.} \quad (6)$$

where A^- is a symmetric generalised inverse, λ_i and δ_j are the nonzero eigenvalues of Σ and Ξ respectively. Let $H = (H_1|H_2) \in \mathcal{O}(N)$ and $P = (P_1|P_2) \in \mathcal{O}(m)$ be matrices associated with

the spectral decomposition of matrices Σ and Ξ respectively with $H_1 \in V_{k,N}$, $H_2 \in V_{N-k,N}$, $P'_1 \in V_{r,m}$ and $P'_2 \in V_{m-r,m}$, see Díaz-García et al.(1997).

Alternatively, this density can be written as

$$dF_X(X) = \frac{1}{(2\pi)^{rk/2} \left(\prod_{i=1}^r \lambda_i^{k/2} \right) \left(\prod_{j=1}^k \delta_j^{r/2} \right)} \exp \left(-\frac{1}{2} \Sigma^- (X - \mu)' \Xi^- (X - \mu) \right) (dX), \quad (7)$$

where (dX) is the Hausdorff measure, which coincides with that of Lebesgue when it is defined on the subspace \mathcal{M} given by the hyperplane (6), see Díaz-García et al.(1997), Cramér [p. 297, 1999] and Billingley [p. 247, 1979]. Explicitly, if $q = \min(r, k)$, (dX) would be given by (1) and/or considering the SVD, (dX) is defined by (3).

Definition 2. [Singular and Nonsingular Wishart and Pseudo-Wishart Distributions.] *Let us suppose that $Y \sim \mathcal{N}_{N \times m}^{k,r}(\mu, \Sigma, \Xi)$, with $r(\Sigma) = r \leq m$, $r(\Xi) = k \leq N$ and let $q = \min(r, k)$, then the density of $S = Y' \Xi^- Y$ is given by*

$$\frac{\pi^{k(q-r)/2} |L|^{(k-m-1)/2}}{2^{kr/2} \Gamma_q \left[\frac{1}{2} k \right] \left(\prod_{i=1}^r \lambda_i^{k/2} \right)} \text{etr} \left(-\frac{1}{2} \Sigma^- S - \frac{1}{2} \Omega \right) {}_0F_1 \left(\frac{1}{2} k; \frac{1}{4} \Omega \Sigma^- S \right) \quad (8)$$

$$P_2 S P'_2 = P_2 \mu' \Xi^- \mu P'_2 \quad a. s. \quad (9)$$

where $S = W_1 L W'_1$, Σ^- is a symmetric generalised inverse of Σ , $\Omega = \Sigma^- \mu' \Xi^- \mu$, Ξ^- is a symmetric generalised inverse of $\Xi = Q' Q$ with Q $k \times N$ matrix, $r(Q) = k$, $P' \Sigma P = \Delta_\Sigma$, and ${}_0F_1(\cdot)$ is a hypergeometric function with a matrix argument (see Díaz-García et al.(1997).

Again, observe that this density can be written as

$$dF_S(S) = \frac{\pi^{k(q-r)/2} |L|^{(k-m-1)/2}}{2^{kr/2} \Gamma_q \left[\frac{1}{2} k \right] \left(\prod_{i=1}^r \lambda_i^{k/2} \right)} \text{etr} \left(-\frac{1}{2} \Sigma^- S - \frac{1}{2} \Omega \right) {}_0F_1 \left(\frac{1}{2} k; \frac{1}{4} \Omega \Sigma^- S \right) (dS), \quad (10)$$

where (dS) is the Hausdorff measure, which coincides with the Lebesgue measure when the latter is defined on the manifold $\mathcal{S}_m^+(q)$. Explicitly, (dS) is given by (2) and/or with spectral factorisation by (4). Moreover, if S has a density of (10), this is denoted by writing $S \sim \mathcal{W}_m(q, k, \Sigma, \Omega)$ if $k \geq r$ ($N \geq m$) for the case of Wishart distribution, and by $S \sim \mathcal{PW}_m(q, k, \Sigma, \Omega)$ if $k < r$ ($N < m$) for the case of Pseudo-Wishart distribution.

Let us now consider the linear transformation $\underset{\sim}{A}: \mathbb{R}^n \rightarrow \mathbb{R}^m$, defined by $\underset{\sim}{A}(x) = Ax$, for $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$. For this transformation to be one to one, it is necessary to restrict its domain to the rank of A' , $R(A') = \{x | A'g = x\}$. Thus, the following definition is obtained:

Definition 3. [More-Penrose inverse.] *If $A \in \mathbb{R}^{m \times n}$, define the linear transformation $\underset{\sim}{A}^+$: $\mathbb{R}^m \rightarrow \mathbb{R}^n$ by $\underset{\sim}{A}^+(x) = 0$ if $x \in R(A)^\perp$ and $\underset{\sim}{A}^+(x) = \left(\underset{\sim}{A} \Big|_{R(A')} \right)^{-1} x$ if $x \in R(A)$. The*

matrix of \tilde{A}^+ is denoted A^+ and called the generalised or More-Penrose inverse of A , see Campbell and Meyer [pp. 8-9, 1979].

Given a matrix A , various methods may be applied to determine its generalised inverse. For example, if $A \in \mathcal{L}_{m,N}^+(q)$ and $A = H_1 D P_1'$ is the nonsingular part of the decomposition into singular values of A , its generalised inverse is given by $A^+ = P_1 D^{-1} H_1'$, see Rao [pp.76-77, 1973].

3. JACOBIANS

In this section, we determine the Jacobian of the transformation $Y = X^+$, both for the singular rectangular case and for that in which X is a non-negative defined matrix.

Theorem 1. *If $W = V^+$ with $V \in \mathcal{S}_m^+(r)$. Then*

$$(dW) = |D_\lambda|^{-2m+r-1} (dV) = \prod_{i=1}^r \lambda_i^{2m+r-1} (dV), \quad (11)$$

where $V = H_1 D_\lambda H_1'$, is the nonsingular part of the spectral decomposition of V , with $D_\lambda = \text{diag}(\lambda_1, \dots, \lambda_r)$, $\lambda_1 > \dots > \lambda_r > 0$ and $H_1 \in \mathcal{V}_{r,m}$.

Proof : Let $V = H_1 D_\lambda H_1'$ be the nonsingular part of the spectral decomposition of V , then $W = V^+ = H_1 D_\lambda^{-1} H_1'$. Then, from Proposition 2, we have

$$(dV) = 2^{-m} \prod_{i=1}^r \lambda_i^{m-r} \prod_{i<j} (\lambda_i - \lambda_j) (H_1' dH_1) \wedge \bigwedge_{i=1}^r d\lambda_i, \quad (12)$$

similarly

$$(dW) = 2^{-m} \prod_{i=1}^r (\lambda_i^{-1})^{m-r} \prod_{i<j} (\lambda_i^{-1} - \lambda_j^{-1}) (H_1' dH_1) \wedge \bigwedge_{i=1}^r d\lambda_i^{-1}. \quad (13)$$

Now, ignoring the sign,

$$\bigwedge_{i=1}^r d\lambda_i^{-1} = \bigwedge_{i=1}^r (-1) \frac{d\lambda_i}{\lambda_i^2} = \prod_{i=1}^r \lambda_i^{-2} \bigwedge_{i=1}^r d\lambda_i,$$

and so

$$(dW) = 2^{-m} \prod_{i=1}^r (\lambda_i^{-(m-r+2)}) \prod_{i<j} (\lambda_i^{-1} - \lambda_j^{-1}) (H_1' dH_1) \wedge \bigwedge_{i=1}^r d\lambda_i,$$

from which

$$(H_1' dH_1) \wedge \bigwedge_{i=1}^r d\lambda_i = 2^m \left[\prod_{i=1}^r (\lambda_i^{-(m-r+2)}) \prod_{i<j} (\lambda_i^{-1} - \lambda_j^{-1}) \right]^{-1} (dW). \quad (14)$$

Finally, the result is obtained by substituting (14) in (12). Note that (ignoring the sign),

$$\prod_{i<j} \frac{(\lambda_i - \lambda_j)}{(\lambda_i^{-1} - \lambda_j^{-1})} = \prod_{i<j} \lambda_i \lambda_j = \prod_{i=1}^r \lambda_i^{r-1}.$$

Remark 1. Note that if $r = m$, that is $V \in \mathcal{S}_m$, then $|D_\lambda| = |V|$, thus obtaining ■

$$(dW) = |V|^{-(m+1)}(dV),$$

see, for example, Muirhead [Theorem 2.1.8, p.59, 1982].

This result can easily be extended to the case of a rectangular matrix of incomplete rank, considering the nonsingular part of the decomposition in singular values instead of the nonsingular part of the spectral decomposition, and thus

Theorem 2. Let $Y = X^+$, with $X \in \mathcal{L}_{m,N}^+(q)$. Then

$$(dY) = |D_\sigma|^{-2(N+m-q)}(dX) = \prod_{i=1}^q \sigma_i^{-2(N+m-q)}(dX), \quad (15)$$

where $X = H_1 D_\sigma P_1'$, is the nonsingular part of the decomposition in singular values of X , with $D_\sigma = \text{diag}(\sigma_1, \dots, \sigma_q)$, $\sigma_1 > \dots > \sigma_q > 0$, $H_1 \in \mathcal{V}_{q,N}$ and $P_1 \in \mathcal{V}_{q,m}$.

Proof : The demonstration is analogous to that given for Theorem 1, but in this case taking into account Proposition 1. ■

4. GENERALISED INVERSE WISHART AND PSEUDO-WISHART DISTRIBUTIONS AND MATRIX-VARIATE SINGULAR T -DISTRIBUTIONS

Given the matrix $S \sim \mathcal{W}_m(q, k, \Sigma, \Omega)$ ($o \sim \mathcal{PW}_m(q, k, \Sigma, \Omega)$), we now wish to determine the distribution of $U = S^+$. This distribution is called m -dimensional generalised inverse Wishart (or Pseudo-Wishart), of rank q , with k degrees of freedom, a scale matrix $\Sigma^- \geq 0$ and noncentrality parameter matrix Ω . Note that, following the notation of Srivastava and Khatri [p. 72, 1979] and/or Díaz-García et al. (1997, when $k \geq r$ ($N \geq m$)) the generalised inverse Wishart distribution (singular or nonsingular) is obtained. Otherwise, the generalised inverse Pseudo-Wishart distribution (singular or nonsingular) is obtained. These cases are denoted, respectively, by $U \sim \mathcal{W}_m^+(q, k + m + 1, \Sigma^-, \Omega)$ and by $U \sim \mathcal{PW}_m^+(q, k + m + 1, \Sigma^-, \Omega)$.

Theorem 3. Let $S \sim \mathcal{W}_m(q, k, \Sigma, \Omega)$ ($o \sim \mathcal{PW}_m(q, k, \Sigma, \Omega)$) and let $U = S^+$. Then the density of U is given by

$$dF_U(U) = \frac{\pi^{k(q-r)/2} |D_T|^{-(k+3m-2q+1)/2}}{2^{kr/2} \Gamma_q \left[\frac{1}{2}k \right] \left(\prod_{i=1}^r \lambda_i^{k/2} \right)} \text{etr} \left(-\frac{1}{2} \Sigma^- U^+ - \frac{1}{2} \Omega \right) {}_0F_1 \left(\frac{1}{2}k; \frac{1}{4} \Omega \Sigma^- U^+ \right) (dU), \quad (16)$$

where $U = H_1 D_T H_1$ is the nonsingular part of the spectral decomposition of U , with $D_T = \text{diag}(t_1, \dots, t_q)$, $t_1 > \dots > t_q > 0$, $H_1 \in \mathcal{V}_{q,m}$ and the measure (dU) is explicitly given by

$$(dU) = 2^{-m} \prod_{i=1}^q t_i^{m-q} \prod_{i < j} (t_i - t_j) (H_1' dH_1) \wedge \bigwedge_{i=1}^q dt_i. \quad (17)$$

Proof : From $U = H_1 D_T H_1$ the proof is immediate from equation (10) and Theorem 1. ■
This distribution and some of its properties have been studied by various authors, for the central Wishart nonsingular distribution (that is when $\Omega = 0$, $m \leq N = k$ and $q = r = m$), see for example Press [Section 5.2, 1982], Box and Tiao [p.460, 1973] and Gupta and Nagar [Section 3.4, 2000]. By adopting the notation of Box and Tiao [p.460, 1973] the following is obtained:

Definition 1. [Generalised Inverse Wishart and Pseudo-Wishart Distributions] *Let $U \in \mathcal{S}_m^+(q)$ be a random matrix. U is said to have a Wishart or Pseudo-Wishart inverse-generalised non-central distribution with ν degrees of freedom, a scale matrix G and a noncentrality parameter matrix Ω , this fact being denoted by $U \sim \mathcal{W}_m^+(q, \nu, G, \Omega)$ and by $U \sim \mathcal{PW}_m^+(q, \nu, G, \Omega)$ respectively, if the density function is given by*

$$dF_U(U) = \frac{\pi^{(\nu+m-1)(q-r)/2} \left(\prod_{i=1}^r \delta_i^{(\nu+m-1)/2} \right)}{2^{(\nu+m-1)r/2} \Gamma_q \left[\frac{1}{2}(\nu+m-1) \right] |D_T|^{(\nu+4m-2q)/2}} \text{etr} \left(-\frac{1}{2} G U^+ - \frac{1}{2} \Omega \right) {}_0F_1 \left(\frac{1}{2}(\nu+m-1); \frac{1}{4} \Omega G U^+ \right) (dU), \quad (18)$$

where $\Omega = G \mu' \Xi^- \mu$, $U = H_1 D_T H_1$ is the nonsingular part of the spectral decomposition of U , with $D_T = \text{diag}(t_1, \dots, t_q)$, $t_1 > \dots > t_q > 0$, $H_1 \in \mathcal{V}_{q,m}$; $G = R_1 D_\delta R_1'$ is the nonsingular part of the spectral decomposition of G , with $D_\delta = \text{diag}(\delta_1, \dots, \delta_q)$, $R_1 \in \mathcal{V}_{q,m}$, and where the measure (dU) is explicitly given by (17).

Note that if $\Omega = 0$ we obtain the central (singular and nonsingular) generalised inverse Wishart and Pseudo-Wishart distributions, whose density is given by

$$dF_U(U) = \frac{\pi^{(\nu+m-1)(q-r)/2} \left(\prod_{i=1}^r \delta_i^{(\nu+m-1)/2} \right)}{2^{(\nu+m-1)r/2} \Gamma_q \left[\frac{1}{2}(\nu+m-1) \right] |D_T|^{(\nu+4m-2q)/2}} \text{etr} \left(-\frac{1}{2} G U^+ \right) (dU) \quad (19)$$

We now wish to determine an expression for the density of a matrix-variate singular T -distribution. When T is nonsingular, both in the central and in the noncentral cases, this distribution has been studied by various authors, including Box and Tiao [Section 8.4.3, 1973], Press [Section 6.2.3, 1982] and Gupta and Nagar [Chapter 4, 2000]. By considering the expression for the density of the matrix-variate T , given by Press [p. 139, 1982] equation (6.2.7) and the notation of Box and Tiao [p. 441, 1972] the following result is obtained:

Theorem 4. *Let $Y \sim \mathcal{MT}_{N \times m}(\mu, \Sigma, \Xi, \nu)$, with Σ $m \times m$, $r(\Sigma) = r < m$ or Ξ $N \times N$, $r(\Xi) = k < N$. This is called the matrix-variate singular T -distribution and is denoted*

$$Y \sim \mathcal{MT}_{N \times m}^{k,r}(\mu, \Sigma, \Xi, \nu)$$

Omitting the supra-indices when $r = m$ and $k = N$. Its density function is given by

$$dF_Y(Y) = \frac{\Gamma_r \left[\frac{1}{2}(\nu + k + r - 1) \right] \prod_{i=1}^r \lambda_i^{(\nu+r-1)/2}}{\pi^{rk/2} \Gamma_r \left[\frac{1}{2}(\nu + r - 1) \right] \prod_{j=1}^k \delta_j^{r/2}} |\Sigma + (Y - \mu)' \Xi^- (Y - \mu)|^{-(\nu+k+r-1)/2} (dY), \quad (20)$$

where A^- is a symmetric generalised inverse, λ_i y δ_j are the non-null eigenvalues of S and Ξ respectively, and (dY) is the Hausdorff measure (see Definition 1).

Proof: The proof is parallel to that given for the normal case, see Díaz-García et al.(1997), or Díaz-García and Gutiérrez-Jáimez (2003). ■

5. SOME APPLICATIONS

In this section, the Bayesian inference is performed on Θ and Σ in the linear model of the singular multivariate full rank model defined by:

$$Y = X \Theta + \epsilon, \quad (21)$$

$n \times m$ $n \times k$ $k \times m$ $n \times m$

where $p(\epsilon|\Theta, \Sigma) \equiv \mathcal{N}_{n \times m}^{n,r}(0, \Sigma \otimes I_n)$ with $\Sigma \geq 0$, $r(\Sigma) = r \leq m < n$ and $r(X) = k$. Let $S(\Theta)$ be the symmetric matrix

$$\begin{aligned} S(\Theta) &= (Y - X\Theta)'(Y - X\Theta) \\ &= (Y - \hat{\Theta})'(Y - \hat{\Theta}) + (\Theta - \hat{\Theta})'X'X(\Theta - \hat{\Theta}) \\ &= A + (\Theta - \hat{\Theta})'X'X(\Theta - \hat{\Theta}) \end{aligned}$$

with $A = (Y - \hat{\Theta})'(Y - \hat{\Theta})$ y $\hat{\Theta} = X^+Y = (X'X)^{-1}X'Y$ is the estimator of least squares of Θ . From these observations, and from expression (7), the likelihood function can be written as

$$\begin{aligned} \mathcal{L}(\Theta, \Sigma|Y) &\propto dP(\epsilon|\Theta, \Sigma)(d\Theta)(d\Sigma) \\ &\propto \prod_{i=1}^r \lambda_i^{-n/2} \text{etr} \left(-\frac{1}{2}\Sigma^+ S(\Theta) \right) (d\Theta)(d\Sigma), \end{aligned} \quad (22)$$

where λ_i , $i = 1, 2, \dots, r$, are the non-null eigenvalues of Σ .

For the joint prior distribution of the parameters (Θ, Σ) , assume that Θ and Σ are approximately independent, that is, $p(\Theta, \Sigma) \doteq p(\Theta)p(\Sigma)$ and Θ is locally uniform, thus $p(\Theta) \propto \text{constant}$. We still have to determine the joint prior distribution of the $mr - r(r-1)/2$ mathematically-independent elements in Σ . For this purpose, and proceeding in an analogous way to Box and Tiao [Section 8.2.2,1972], for Theorem 1 we find that

$$dP(\Sigma) \propto \prod_{i=1}^r \lambda_i^{-(2m-r+1)/2} (d\Sigma).$$

From which

$$dP(\Theta, \Sigma) \propto \prod_{i=1}^r \lambda_i^{-(2m-r+1)/2} (d\Theta)(d\Sigma). \quad (23)$$

Thus, the joint posterior distribution of the parameters (Θ, Σ) is

$$\begin{aligned} dP(\Theta, \Sigma|Y) &\propto \mathcal{L}(\Theta, \Sigma|Y)p(\Theta, \Sigma)(d\Theta)(d\Sigma) \\ &\propto \prod_{i=1}^r \lambda_i^{-(n+2m-r+1)/2} \text{etr}\left(-\frac{1}{2}\Sigma^+ S(\Theta)\right) (d\Theta)(d\Sigma). \end{aligned} \quad (24)$$

By applying (19) we can integrate (24) with respect to $\Sigma \in \mathcal{S}_m^+(r)$. Thus, assuming that $S(\Theta) > 0$ we find that the marginal posterior distribution of Θ is

$$dP(\Theta|Y) \propto |S(\Theta)|^{-(n-m+r)/2} \quad (25)$$

Analogously, by integrating (24), now with respect to $\Theta \in \mathcal{L}_{m,k}^+(r)$, and by applying (20), we find that the marginal posterior distribution of Σ is

$$dP(\Sigma|Y) \propto \prod_{i=1}^r \lambda_i^{-(n+2m-r-k+1)/2} \text{etr}\left(-\frac{1}{2}\Sigma^+ A\right) (d\Sigma) \quad (26)$$

It is also possible to determine the conditional posterior distribution of Θ given Σ , which is given by

$$dP(\Theta|\Sigma, Y) = \frac{p(\Theta, \Sigma|Y)}{p(\Sigma|Y)}(d\Theta) \propto \frac{\text{etr}\left(-\frac{1}{2}\Sigma^+ S(\Theta)\right)}{\text{etr}\left(-\frac{1}{2}\Sigma^+ A\right)}(d\Theta),$$

from which we obtain that

$$dP(\Theta|\Sigma, Y) \propto \text{etr}\left(-\frac{1}{2}\Sigma^+ (\Theta - \hat{\Theta})' X' X (\Theta - \hat{\Theta})\right) (d\Theta) \quad (27)$$

Finally, it is possible to find the conditional posterior distribution of Σ given Θ as

$$dP(\Sigma|\Theta, Y) = \frac{p(\Theta, \Sigma|Y)}{p(\Theta|Y)}(d\Sigma) \propto \frac{\prod_{i=1}^r \lambda_i^{-(n+2m-r+1)/2} \text{etr}\left(-\frac{1}{2}\Sigma^+ S(\Theta)\right)}{|S(\Theta)|^{-(n-m+r)/2}}(d\Sigma)$$

from which

$$dP(\Sigma|\Theta, Y) \propto \prod_{i=1}^r \lambda_i^{-(n+2m-r+1)/2} \text{etr}\left(-\frac{1}{2}\Sigma^+ S(\Theta)\right) \quad (28)$$

In summary, the following result is obtained.

Theorem 5. *Given the general multivariate linear model (21), and assuming a noninformative prior joint distribution for the parameters (Θ, Σ) , the following holds:*

(i) *The joint density function of (Θ, Σ) is given by*

$$\begin{aligned} dP(\Theta, \Sigma|Y) = & \frac{|X'X|^{r/2} \prod_{j=1}^r \delta_j^{(\nu+2r-m-1)/2}}{2^{(\nu+k+2r-m-1)m/2} \pi^{kr/2} \Gamma_r\left[\frac{1}{2}(\nu+2r-m-1)\right] \prod_{i=1}^r \lambda_i^{(\nu+k+2m)/2}} \\ & \text{etr}\left(-\frac{1}{2}\Sigma^+ \left(A + (\Theta - \hat{\Theta})' X' X (\Theta - \hat{\Theta})\right)\right) (d\Theta)(d\Sigma), \end{aligned}$$

where $\delta_j, j = 1, 2, \dots, r$, are the non-null eigenvalues of A .

(ii) $\Theta|Y \sim \mathcal{MT}_{k \times m}^{k,r}(\hat{\Theta}, (X'X)^{-1}, A, \nu)$, (iii) $\Theta|\Sigma, Y \sim \mathcal{N}_{k \times m}^{k,r}(\hat{\Theta}, \Sigma, (X'X)^{-1})$, (iv) $\Sigma|Y \sim \mathcal{W}_m^+(r, \nu, A)$ and (v) $\Sigma|\Theta, Y \sim \mathcal{W}_m^+(r, \nu + k, S(\Theta))$, with $\nu = n - (k + r) + 1$.

Remark 2. The above results can easily be extended to the case of the singular multivariate linear model of non-full rank, that is, when $r(X) = h < k$. In general terms, it is sufficient to make the following changes to the previous results; k for h , $(X'X)^{-1}$ for $(X'X)^+$ and $|X'X|$ for $\prod_{i=1}^h \alpha_i$, where α_i are the non-null eigenvalues of $X'X$. Moreover, in this case, $\hat{\Theta}$ is any solution to the system of normal equations, $X'X\hat{\Theta} = X'Y$, which in particular can be followed by considering $\hat{\Theta} = X^+Y$.

6. ACKNOWLEDGMENT

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

REFERENCES

- [1] Billingsley, P. (1986). *Probability and Measure*, 2nd Edition, John Wiley & Sons, New York.
- [2] Box, G. E., and Tiao, G. C. (1972). *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Company, Reading.
- [3] Campbell, S. L., and Meyer, Jr, C. D. (1979). *Generalized Inverses of Linear Transformations*, Pitman, London.
- [4] Cramér, H. (1999). *Mathematical Methods of Statistics*, Nineteenth printing, Princeton University Press, Princeton.
- [5] Díaz-García, J. A., Gutiérrez-Jáimez, R., and Mardia, K. V. (1997). “Wishart and Pseudo-Wishart distributions and some applications to shape theory”, *J. Multivariate Anal.* **63**, 73-87.
- [6] Díaz-García, J. A. and Gutiérrez, J. R. (1997). “Proof of the conjectures of H. Uhlig on the singular multivariate beta and the jacobian of a certain matrix transformation”, *Ann. Statist.* **25**, 2018-2023.
- [7] Díaz-García J. A. and Gutiérrez-Jáimez R. (2003). “Wishart and Pseudo-Wishart distributions and some applications to shape theory II”, *Comunicación Técnica* No. I-03-01 (PE/CIMAT), <http://www.cimat.mx/biblioteca/RepTec>.
- [8] Eaton, M. L. (1983). *Multivariate Statistics: A Vector Space Approach*, John Wiley & Sons., New York.
- [9] Farrell, R. M. (1985). *Multivariate Calculus: Use of the Continuous Groups*, Springer, New York.

- [10] Golub, G. H., and Van Loan, C. F. (1996). *Matrix Computations*, Johns Hopkins, Baltimore and London.
- [11] Gupta, A. K., and Nagar, D. K. (2000). *Matrix Variate Distributions*, Chapman & Hall/CRC, Washington, D. C.
- [12] James, A. T. (1954). "Normal multivariate analysis and the orthogonal group", *Ann. Math. Statist.* **25**, 40-75.
- [13] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, New York.
- [14] Press, S. J. (1982). *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference*, Second Edition, Robert E. Krieger Publishing Company, Malabar, Florida.
- [15] Rao, C. R. (1973). *Linear Statistical Inference and its Applications* (2nd ed.), John Wiley & Sons, New York.
- [16] Srivastava, M. S., and Khatri, C. G. (1979). *An Introduction to Multivariate Analysis*, North-Holland Publ., Amsterdam.
- [17] Uhlig, H. (1994). "On singular Wishart and singular multivariate beta distributions", *Ann. Statistic.* **22**395-405.
- [18] Xu, J. L. (1990). "Inverse Dirichlet distribution and its applications", in, Fang, K. T., and Anderson, T. W. (1990). *Statistical Inference in Elliptically Contoured and Related Distributions*, Allerton Press Inc., New York, 103-113.