The Annals of Statistics
1994, Vol. 22, No. 1, 395–405

ON SINGULAR WISHART AND SINGULAR MULTIVARIATE BETA DISTRIBUTIONS

BY HARALD UHLIG

Princeton University

This paper extends the study of Wishart and multivariate beta distributions to the singular case, where the rank is below the dimensionality. The usual conjugacy is extended to this case. A volume element on the space of positive semidefinite $m \times m$ matrices of rank $n < m$ is introduced and some transformation properties established. The density function is found for all rank-n Wishart distributions as well as the rank-1 multivariate beta distribution. To do that, the Jacobian for the transformation to the singular value decomposition of general $m \times n$ matrices is calculated. The results in this paper are useful in particular for updating a Bayesian posterior when tracking a time-varying variance–covariance matrix.

1. Introduction. Consider n draws $Y_j, j = 1, \ldots, n$, from a normal distribution $\mathcal{N}(0, \Sigma)$, where Σ is $m \times m$ and positive definite. The random variable $X = \sum_{j=1}^{n} Y_j Y_j'$ has a Wishart distribution $\mathcal{W}_m(n, \Sigma)$. Usually, Wishart distributions are studied only for $n > m-1$. This paper extends the study of Wishart as well as multivariate beta distributions to the singular case, where $0 < n < m$, n an integer, that is, where the rank of the random matrix is below its dimensionality. The usual conjugacy between Wishart and beta distributions [see Muirhead (1982), Theorem 3.3.1] is extended to this case (see Theorems 1 and 7). A volume element on the space of positive definite $m \times m$ matrices of rank $n < m$ is introduced (see Theorem 2) and some transformation properties established (see Theorems 3 and 4)—this is necessary in order to talk sensibly about densities on that space. The density is found for the rank-n Wishart distribution for all integers n, $0 < n < m$ (see Theorem 6) and the rank-1 multivariate beta distribution (see Theorem 7). To do that, the Jacobian for the transformation to the singular value decomposition of general $m \times n$ matrices is calculated (see Theorem 5). This paper thus extends the results established by Fisher (1915), Wishart (1928), James (1954), Khatri (1959), Olkin and Roy (1954) and Olkin and Rubin (1964). Their results are presented clearly and concisely in book form in Muirhead (1982): We will follow his terminology and formulations closely.

The results in this paper are, in particular, fundamental and useful for updating a Bayesian posterior when tracking a time-varying variance–covariance matrix: that, in fact, motivated this investigation [see Uhlig (1992)]. Imagine the following time series model, which is a simple multivariate state space alternative to the popular ARCH models. [For an overview of the extensive

Received July 1992; revised June 1993.

AMS 1991 *subject classifications.* Primary 62H10; secondary 62E15.

Key words and phrases. Wishart distribution, beta distribution, Stiefel manifold, singular matrix distributions, conjugacy.

ARCH literature, see Bollerslev, Chou and Kroner (1992). For a univariate state space specification, see Shephard (1994).] There is an unobservable precision matrix P_t, evolving over time according to

$$P_t = \frac{p+1}{p}\mathcal{U}(P_{t-1})' Q_t \mathcal{U}(P_{t-1}), \tag{1}$$

where $\mathcal{U}(P_{t-1})$ is the upper-triangular Cholesky factor, that is, that upper-triangular matrix T with positive diagonal elements which satisfies $P_{t-1} = T'T$. Suppose that the Q_t are drawn i.i.d. from a multivariate beta distribution $\mathcal{B}_m(p/2, 1/2)$. Suppose further, a researcher starts with the prior that P_{t-1} is Wishart distributed, $P_{t-1} \sim \mathcal{W}_m(\lambda^{-1}, \lambda S_{t-1}^{-1})$, where $\lambda = 1/(p+1)$, so that $E[P_{t-1}]^{-1} = S_{t-1}$. If the usual conjugacy between Wishart and multivariate beta distributions holds, then the prior for P_t is a Wishart $\mathcal{W}_m(p, S_{t-1}^{-1}/p)$. Suppose now that the researcher observes a single draw Y_t from a multivariate normal distribution with that precision matrix, $Y_t \sim \mathcal{N}(0, P_t^{-1})$. The posterior for P_t is then given by a Wishart $\mathcal{W}_m(\lambda^{-1}, \lambda S_t^{-1})$, where $S_t = \lambda Y_t Y_t' + (1-\lambda) S_{t-1}$ so that $E[P_t]^{-1} = S_t$ and the game can begin anew. To find the parameter p governing the degree of time variation, the explicit likelihood function is needed. The problem with these arguments is that the singular multivariate beta distributions $\mathcal{B}_m(p/2, 1/2)$ have yet to be defined and the "usual conjugacy" between Wishart and this multivariate beta distribution has yet to be established. To do that, singular Wishart distributions have to be analyzed as well since they are fundamental for the study of singular multivariate beta distributions. Furthermore, in order to state the likelihood function explicitly, the density function for a $\mathcal{B}_m(p/2, 1/2)$-distributed random variable Q has yet to be found, since $I_m - Q$ is of rank 1 and thus is singular almost surely. Solving these problems is the purpose of this paper. In the course of doing so, some generally useful theorems for the analysis of multivariate random variables are established.

2. Results. Unless stated otherwise, all our notation, definitions and terminology follow Muirhead (1982). First, we generalize the definition of multivariate beta distribution $\mathcal{B}_m(n/2, p/2)$ to integers $0 < n < m$. Let m and n be positive integers, $p > m-1$ and Σ be of size $m \times m$ and positive definite. Recall Definition 3.1.3 in Muirhead (1982), that a random variable A is $\mathcal{W}_m(n, \Sigma)$-distributed if A can be written as

$$A = \Sigma_{j=1}^{n} Y_j Y_j', \quad \text{with } Y_j \sim \mathcal{N}(0, \Sigma) \text{ i.i.d.}$$

For a positive definite matrix S, let $\mathcal{U}(S)$ denote the upper-triangular Cholesky factor, that is, that upper-triangular matrix T with positive diagonal elements which satisfies $S = T'T$.

DEFINITION 1. A random variable X is $\mathcal{B}_m(n/2, p/2)$-distributed, if it can be written as

$$X = \mathcal{U}(A+B)'^{-1} A \mathcal{U}(A+B)^{-1},$$

where $A \sim \mathcal{W}_m(n, I_m)$ and $B \sim \mathcal{W}_m(p, I_m)$ with A represented as above for $\Sigma = I$ and the Y_j, $j = 1, \dots, n$, independent from B. X is $\mathcal{B}_m(n/2, p/2)$-distributed if $A \sim \mathcal{W}_m(p, I_m)$ and $B \sim \mathcal{W}_m(n, I_m)$ instead.

This definition is molded after Theorem 3.3.1 in Muirhead (1982) and is therefore contained as a special case in Definition 3.3.2 in Muirhead (1982), for $n > m - 1$. However, for $n \leq m - 1$, this definition is new and Theorem 3.3.1 in Muirhead (1982) needs to be established for these parameters as well. This is done in the following theorem, which is stated "backwards" from the version in Muirhead (1982) to make it particularly suitable for the purpose of posterior updating alluded to in the Introduction.

THEOREM 1. *Let m and n be positive integers and let $p > m - 1$. Let $H \sim \mathcal{W}_m(p + n, \Sigma)$ and $Q \sim \mathcal{B}_m(p/2, n/2)$ be independent. Then*

$$G \equiv \mathcal{U}(H)' Q \mathcal{U}(H) \sim \mathcal{W}_m(p, \Sigma).$$

PROOF. The theorem follows from the following somewhat broader claim:

CLAIM. Let $A \sim \mathcal{W}_m(p, I_m)$, $B = \sum_{j=1}^n Y_j Y_j'$, with $Y_j \sim \mathcal{N}(0, I_m)$ i.i.d., and $H \sim \mathcal{W}_m(p + n, \Sigma)$, where A, Y_j, $j = 1, \dots, n$, and H are independent. Define $C \equiv A + B$, $Q \equiv \mathcal{U}(C)'^{-1} A \mathcal{U}(C)^{-1}$, $G \equiv \mathcal{U}(H)' Q \mathcal{U}(H)$ and $D \equiv H - G$. Then $C \sim \mathcal{W}_m(p + n, I_m)$, $G \sim \mathcal{W}_m(p, \Sigma)$ and $D = \sum_{j=1}^n Z_j Z_j'$ with $Z_j \sim \mathcal{N}(0, \Sigma)$, where C, G and Z_j, $j = 1, \dots, n$, are independent.

The proof mimics the proof of Muirhead [(1982), Theorem 3.3.1]. Define $Z_j = \mathcal{U}(H)' \mathcal{U}(C)'^{-1} Y_j$ and note that $D = \sum_{j=1}^n Z_j Z_j'$. It follows from Muirhead [(1982), Theorem 2.1.4] that

$$\begin{aligned}
&(dA) \wedge (dH) \wedge (dY_1) \wedge \cdots \wedge (dY_n) \\
&\quad = (dC) \wedge (dH) \wedge (dY_1) \wedge \cdots \wedge (dY_n) \\
&\quad = (\det H)^{-n/2} (\det C)^{n/2} (dC) \wedge (dH) \wedge (dZ_1) \wedge \cdots \wedge (dZ_n) \\
&\quad = (\det H)^{-n/2} (\det C)^{n/2} (dC) \wedge (dG) \wedge (dZ_1) \wedge \cdots \wedge (dZ_n),
\end{aligned}$$

exploiting $G = H - D$ for the last equality. Writing out the densities, it now follows that

$$\begin{aligned}
&\left(2^{mp/2} \Gamma_m(p/2)\right)^{-1} \operatorname{etr}(-A/2) (\det A)^{(p-m-1)/2} \\
&\quad \times \left(2^{m(n+p)/2} \Gamma_m((n+p)/2) (\det \Sigma)^{(n+p)/2}\right)^{-1} \operatorname{etr}(-\Sigma^{-1} H/2) (\det H)^{(n+p-m-1)/2} \\
&\quad \times (2\pi)^{-mn/2} \operatorname{etr}(-B/2)\, (dA) \wedge (dH) \wedge (dY_1) \wedge \cdots \wedge (dY_n) \\
&\quad = \left(2^{m(n+p)/2} \Gamma_m((n+p)/2)\right)^{-1} \operatorname{etr}(-C/2) (\det C)^{(n+p-m-1)/2} \\
&\qquad \times \left(2^{mp/2} \Gamma_m(p/2) (\det \Sigma)^{p/2}\right)^{-1} \operatorname{etr}(-\Sigma^{-1} G/2) (\det G)^{(p-m-1)/2} \\
&\qquad \times (2\pi)^{-mn/2} (\det \Sigma)^{-n/2} \operatorname{etr}(-D/2)\, (dC) \wedge (dG) \wedge (dZ_1) \wedge \cdots \wedge (dZ_n),
\end{aligned}$$

exploiting $\det A = \det C \det Q$, $H = G + D$ and $\det H = \det G / \det Q$. Inspecting the latter density finishes the proof. □

Let $m > n > 0$ be integers. For the computation of a likelihood function, say, a density on a space of appropriate dimensionality is needed. Densities do not exist for $V \sim \mathcal{W}_m(n, \Sigma)$ or $X \sim \mathcal{B}_m(p, n/2)$ on the space of symmetric $m \times m$ matrices, since V and $I_m - X$ are singular and of rank n almost surely [see Muirhead (1982), Theorem 3.1.4]. As shown below, however, densities do exist on the $(mn - n(n-1)/2)$-dimensional manifold of rank-n positive semidefinite $m \times m$ matrices S with n distinct positive eigenvalues; denote that manifold by $\mathcal{S}^+_{m,n}$. A natural global coordinate system for this manifold is to use the decomposition $S = H_1 L H_1'$, where L is $n \times n$, diagonal, $L = \operatorname{diag}(l_1, \ldots, l_n)$ with $l_1 > l_2 > \cdots > 0$ and where $H_1 \in V_{n,m}$, the $(mn - n(n+1)/2)$-dimensional Stiefel manifold of $m \times n$ matrices H_1 with orthonormal columns, $H_1' H_1 = I_n$. This parameterization is unique up to the assignment of n arbitrary signs to the columns of H_1. The task is to define the volume element (dS): with the chosen parameterization, (dS) needs to be defined as some function of H_1 and L multiplied with $(H_1'\,dH_1) \wedge \bigwedge_{i=1}^n dl_i$. [We follow Muirhead (1982), page 56, in ignoring signs of the overall differential and defining only positive integrals. For the definition of $(H_1'\,dH_1)$, see Muirhead (1982), page 63 and the discussion following page 67.] Note that $dS = dH_1 L H_1' + H_1\,dL\,H_1' + H_1 L\,dH_1'$.

Find an $m \times (m-n)$ matrix H_2, so that $H \equiv [H_1 \vdots H_2] \in O(m)$, that is, so that $H'H = I_m$, and let $\mathbf{h}_i$ be the i-th column of H. Let $R = H'\,dS\,H$ and calculate that

$$R = H'\,dS\,H = \begin{bmatrix} R_a & R_b' \\ R_b & \mathbf{0} \end{bmatrix},$$

where

$$\begin{aligned} R_a &= H_1'\,dH_1\,L + dL + (H_1'\,dH_1\,L)' \\ &= \begin{bmatrix} dl_1 & (l_1 - l_2)\mathbf{h}_2'\,d\mathbf{h}_1 & \cdots & (l_1 - l_n)\mathbf{h}_n'\,d\mathbf{h}_1 \\ (l_1 - l_2)\mathbf{h}_2'\,d\mathbf{h}_1 & dl_2 & \cdots & (l_2 - l_n)\mathbf{h}_n'\,d\mathbf{h}_2 \\ \vdots & \vdots & \ddots & \vdots \\ (l_1 - l_n)\mathbf{h}_n'\,d\mathbf{h}_1 & (l_2 - l_n)\mathbf{h}_n'\,d\mathbf{h}_2 & \cdots & dl_n \end{bmatrix} \end{aligned}$$

[exploiting the skew symmetry of $H_1'\,dH_1$, see Muirhead (1982), bottom of page 64] and

$$\begin{aligned} R_b &= H_2'\,dH_1\,L \\ &= \begin{bmatrix} l_1 \mathbf{h}_{n+1}'\,d\mathbf{h}_1 & \cdots & l_n \mathbf{h}_{n+1}'\,d\mathbf{h}_n \\ \vdots & \ddots & \cdots \\ l_1 \mathbf{h}_m'\,d\mathbf{h}_1 & \cdots & l_n \mathbf{h}_m'\,d\mathbf{h}_n \end{bmatrix}. \end{aligned}$$

Appealing to the analogy that (dS) would be the exterior product of all distinct, nonzero entries in R by Theorem 2.1.6 in Muirhead (1982), if S were a full-rank symmetric matrix, we define (dS) to be the exterior product of all entries on and below the main diagonal of R_a and of all entries in R_b times a factor 2^{-n} to correct for the fact that each matrix S is the image of 2^n decompositions H_1LH_1' due to the arbitrary assignment of signs to the columns of H_1. We therefore get the following theorem.

THEOREM 2. *Let $m > n > 0$ be integers. On the space $\mathcal{S}_{m,n}^{+}$ of positive semidefinite $m \times m$ matrices S of rank n with n distinct positive eigenvalues, the volume element (dS) is*

$$(dS) \equiv 2^{-n} \prod_{i=1}^{n} l_i^{m-n} \prod_{i<j}^{n} (l_i - l_j)(H_1' dH_1) \wedge \bigwedge_{i=1}^{n} dl_i, \tag{2}$$

with S represented as $S = H_1LH_1'$, $H_1 \in V_{n,m}$ and $L = \operatorname{diag}(l_1, \dots, l_n)$, $l_1 > l_2 > \cdots > 0$.

We aim at expressing the densities with respect to this volume element subject to the restriction that the density is the same whenever two pairs (H_1, L) and $(\widetilde{H}_1, \widetilde{L})$ differ only in the assignment of signs to the columns of H_1 and $\widetilde{H}_1$. The following useful theorem also justifies defining (dS) in terms of the nonzero and distinct entries of R.

THEOREM 3. *Let $X, Y \in \mathcal{S}_{m,n}^{+}$ be related by $X = QYQ'$, where $Q \in O(m)$. Then $(dX) = (dY)$, where (dX) and (dY) are the volume elements on $\mathcal{S}_{m,n}^{+}$ defined above.*

PROOF. Decompose $Y = H_1LH_1'$, $H_1 \in V_{n,m}$ and $L = \operatorname{diag}(l_1, \dots, l_n), l_1 > l_2 > \cdots > 0$. Note that $G_1 \equiv QH_1 \in V_{n,m}$, so that $X = G_1LG_1'$ is the decomposition for X. Since $(H_1' dH_1)$ is invariant to multiplication on the left with an orthonormal $m \times m$ matrix, that is, since $(G_1' dG_1) = (H_1' dH_1)$ [see Muirhead (1982), bottom of page 69], the result follows. □

The following theorem is a version of Muirhead's Theorem 2.1.6 [Muirhead (1982)] for the case $n = 1$. The general rank-n case is an open problem.

THEOREM 4. *Let $X, Y \in \mathcal{S}_{m,1}^{+}$ be related by $X = BYB'$, where B is $m \times m$ and of full rank. Form the representations $X = G_1KG_1'$, $Y = H_1LH_1'$, where $G_1, H_1 \in V_{1,m}$ and $K, L \in \mathbb{R}$. Then*

$$(dX) = |G_1'BH_1|^m \det(B)(dY).$$

Since $Xv = (KG_1'v)G_1 = (LH_1'B'v)BH_1$, for any $v \in \mathbb{R}^m$, and since $\|G_1\| = 1$ (where we use $\|\cdot\|$ to denote the norm of the $m \times 1$ vector constituting G_1), we have $G_1 = BH_1/\|BH_1\|$ and $K = \|BH_1\|^2 L$, and thus

$$G_1' BH_1 = \| BH_1 \| = (K/L)^{1/2} = 1/ \| B^{-1}G_1 \| , \tag{3}$$

enabling explicit calculation of the expression in the theorem. We conjecture that the formula for the general rank-n case is given by

$$(dX) = \det(G_1'BH_1)^{m+1-n} \det(B)^n (dY).$$

PROOF OF THEOREM 4. We first show this for the case that $B = D$ is diagonal. Find $m \times (m-1)$ matrices G_2 and H_2, so that $G \equiv [G_1 \vdots G_2]$ and $H \equiv [H_1 \vdots H_2]$ satisfy $G'G = I_m$ and $H'H = I_m$, that is, $G, H \in O(m)$. Let g_i and h_i denote the ith columns of G and H. Let $E = G'DH$ and note that

$$G'\, dX\, G = E\, H'\, dY\, H\, E',$$

where, for example,

$$H'\, dY\, H = \begin{bmatrix} dL & (H_2'\, dh_1\, L)' \\ H_2'\, dh_1\, L & \mathbf{0} \end{bmatrix}.$$

Thus, the first column of $E\, H'\, dY\, H\, E'$ has

$$e_{11} e_{i1}\, dL + \sum_{j=2}^{m} (e_{11} e_{ij} + e_{i1} e_{1j}) h_j'\, dh_1\, L$$

as its ith entry. Taking the exterior product over all these entries, ignoring the overall sign for now and using the abbreviation $f_{i1} = e_{11} e_{i1}$, $f_{ij} = e_{11} e_{ij} + e_{i1} e_{1j}$ yields

$$\begin{aligned} (dX) &= \bigwedge_{i=1}^{m} \left(f_{i1}\, dL + \sum_{j=2}^{m} f_{ij} h_j'\, dh_1\, L \right) \\ &= \left(\sum_{\sigma \in \Pi(m)} \operatorname{sgn}(\sigma) \prod_{i=1}^{m} f_{i\sigma(i)} \right) L^{m-1}\, dL \wedge \bigwedge_{j=2}^{m} h_j'\, dh_1 \\ &= (\det F)(dY), \end{aligned}$$

using the skew symmetry of the operator $\wedge$, where $\Pi(m)$ is the set of all permutations of $(1, \ldots, m)$ [cf. Horn and Johnson (1985), page 8] and where F is the matrix $[f_{ij}]_{i=1,j=1}^{m,m}$.

The first column of F is the first column of E, multiplied with $e_{11} = G_1'DH_1 \neq 0$. The jth column of F is the sum of the jth column of E multiplied with e_{11} and the first column of F multiplied with e_{1j}/e_{11}. By the rules about calculating with determinants, it follows that

$$\det F = \det(e_{11} E) = e_{11}^m \det E = e_{11}^m \det D.$$

Since (dX) has a positive sign, the absolute value of e_{11} needs to be taken, demonstrating the claim for $B = D$ diagonal.

For general B, write B as $B = P'DQ$, where $P, Q \in O(m)$ and D is diagonal [see Theorem A9.10 in Muirhead (1982), page 593]. Let $\widetilde{G}_1 = PG_1$, $\widetilde{H}_1 = QH_1$, $\widetilde{X} = PXP'$ and $\widetilde{Y} = QYQ'$. With the aid of the previous theorem and the proof above for diagonal matrices D, it follows that

$$(dX) = (d\widetilde{X}) = (\widetilde{G}_1'D\widetilde{H}_1)^m(\det D)(d\widetilde{Y}) = (G_1'BH_1)^m(\det B)(dY),$$

as claimed. □

The following theorem is an extremely useful cousin of Muirhead's Theorem 2.1.13 [Muirhead (1982), page 63]. The proof is not a straightforward generalization of the proof of that theorem, but it proceeds along similar lines. Let Z be an $m \times n$ $(m \geq n)$ matrix of rank n and with distinct eigenvalues of $Z'Z$. Using the nonsingular part of the singular value decomposition, write $Z = H_1DP'$, where $H_1 \in V_{n,m}$, D is diagonal with $D_{11} > D_{22} > \cdots > D_{nn} > 0$ and $P \in O(n)$: this decomposition is unique up to the arbitrary assignment of signs to columns of P as can be seen upon examination of, for example, Theorem 7.3.5 and its proof in Horn and Johnson [(1985), page 414, with $A = Z'$ there].

THEOREM 5. *Let Z be an $m \times n$ matrix and $Z = H_1DP'$ the nonsingular part of the singular value decomposition, where $H_1 \in V_{n,m}$, D is diagonal with $D_{11} > D_{22} > \cdots > D_{nn} > 0$ and $P \in O(n)$. Then*

$$(dZ) = 2^{-n}(\det D)^{m-n}\prod_{i<j}(D_{ii}^2 - D_{jj}^2)(H_1'\,dH_1) \wedge (dD) \wedge (P'\,dP), \tag{4}$$

where

$$(dD) = \bigwedge_{i=1}^{n} dD_{ii}. \tag{5}$$

PROOF. Find an $m \times (m-n)$ matrix H_2 so that $H \equiv [H_1 \vdots H_2] \in O(m)$. Since

$$dZ = dH_1\,DP' + H_1\,dD\,P' + H_1D\,dP',$$

it follows that

$$H'\,dZ\,P = \begin{bmatrix} H_1'\,dH_1\,D + dD + (P'\,dP\,D)' \\ H_2'\,dH_1\,D \end{bmatrix}. \tag{6}$$

Since $(dZ) = (H'\,dZ\,P)$ by Theorem 2.1.4 in Muirhead (1982), calculating the exterior product of the differential forms on the right-hand side of equation

(6) delivers the solution. The exterior product of all elements in the bottom part $H_2' dH_1 D$ is

$$(H_2' dH_1 D) = (\det D)^{m-n} \bigwedge_{i=1}^{n} \bigwedge_{j=n+1}^{m} \mathbf{h}_j' d\mathbf{h}_i, \tag{7}$$

using Theorem 2.1.1 in Muirhead (1982). The top part

$$T = H_1' dH_1 D + dD + (P' dPD)'$$

is an $n \times n$ matrix of differential forms with entries

$$T_{ij} = \begin{cases} dD_{ii}, & \text{if } i = j, \\ \mathbf{h}_i' d\mathbf{h}_j D_{jj} - D_{ii}\mathbf{p}_i' d\mathbf{p}_j, & \text{if } i > j, \\ \mathbf{h}_j' d\mathbf{h}_i D_{jj} - D_{ii}\mathbf{p}_j' d\mathbf{p}_i, & \text{if } i < j, \end{cases}$$

exploiting the skew symmetry of $H_1' dH_1$ and $P' dP$. To calculate the exterior product of these elements, write that product conveniently as

$$(T) = \bigwedge_{i=1}^{n} T_{ii} \wedge \bigwedge_{i<j} T_{ij} \wedge T_{ji} \tag{8}$$

[where the overall sign of (T) is, as usual, ignored for now]. Examine the entry $T_{ij} \wedge T_{ji}$ in equation (8). Written explicitly, we have

$$\begin{aligned} T_{ij} \wedge T_{ji} &= (\mathbf{h}_j' d\mathbf{h}_i D_{jj} - D_{ii}\mathbf{p}_j' d\mathbf{p}_i) \wedge (\mathbf{h}_j' d\mathbf{h}_i D_{ii} - D_{jj}\mathbf{p}_j' d\mathbf{p}_i) \\ &= -D_{ii}^2 \mathbf{p}_j' d\mathbf{p}_i \wedge \mathbf{h}_j' d\mathbf{h}_i - D_{jj}^2 \mathbf{h}_j' d\mathbf{h}_i \wedge \mathbf{p}_j' d\mathbf{p}_i \\ &\quad + D_{ii}D_{jj}\mathbf{h}_j' d\mathbf{h}_i \wedge \mathbf{h}_j' d\mathbf{h}_i + D_{ii}D_{jj}\mathbf{p}_j' d\mathbf{p}_i \wedge \mathbf{p}_j' d\mathbf{p}_i. \end{aligned}$$

Note that, for example,

$$\begin{aligned} \mathbf{h}_j' d\mathbf{h}_i \wedge \mathbf{h}_j' d\mathbf{h}_i &= \sum_{k<l} H_{kj} H_{lj} (dH_{ki} \wedge dH_{li} + dH_{li} \wedge dH_{ki}) \\ &= 0 \end{aligned}$$

due to the skew symmetry of the operator $\wedge$. Likewise, $-\mathbf{p}_j' d\mathbf{p}_i \wedge \mathbf{h}_j' d\mathbf{h}_i = \mathbf{h}_j' d\mathbf{h}_i \wedge \mathbf{p}_j' d\mathbf{p}_i$, so that

$$T_{ij} \wedge T_{ji} = (D_{ii}^2 - D_{jj}^2)\mathbf{h}_j' d\mathbf{h}_i \wedge \mathbf{p}_j' d\mathbf{p}_i.$$

Combining equations (7) and (8), the overall exterior product of the elements of the right-hand side of equation (6) after appropriate reordering (and again ignoring the overall sign) is thus

$$\begin{aligned} &(H_2' dH_1 D) \wedge (T) \\ &= (\det D)^{m-n} \bigwedge_{i=1}^{n} \bigwedge_{j=n+1}^{m} \mathbf{h}_j' d\mathbf{h}_i \prod_{i<j} (D_{ii}^2 - D_{jj}^2) \bigwedge_{i=1}^{n} dD_{ii} \wedge \bigwedge_{i=1}^{n} \bigwedge_{j=i+1}^{n} \mathbf{h}_j' d\mathbf{h}_i \wedge \bigwedge_{i=1}^{n} \bigwedge_{j=i+1}^{n} \mathbf{p}_j' d\mathbf{p}_i \\ &= (\det D)^{m-n} \prod_{i<j} (D_{ii}^2 - D_{jj}^2) (H_1' dH_1) \wedge (dD) \wedge (P' dP). \end{aligned}$$

Allowing finally for arbitrary assignment of signs to the columns of P makes Z the image of 2^n decompositions $Z = H_1DP'$, so that the density has to be divided by that number in order for the integration over the entire space $P \in O(n)$ to yield the correct result. □

THEOREM 6. *Let $m > n > 0$ be integers. The density for a $\mathcal{W}_m(n, \Sigma)$-distribution on the space $S^+_{m,n}$ of rank-n positive semidefinite $m \times m$ matrices with distinct positive eigenvalues with respect to the volume element (dS) defined above is given by*

$$\frac{\pi^{(-mn+n^2)/2}}{2^{mn/2}\Gamma_n(n/2)(\det \Sigma)^{n/2}} \operatorname{etr}(-\Sigma^{-1}S/2)(\det L)^{(n-m-1)/2}, \tag{9}$$

where $L = \operatorname{diag}(l_1, \dots, l_n)$, $S = H_1LH_1'$.

PROOF. Let $Y = [Y_1 \ \dots \ Y_n]$, where $Y_i \sim \mathcal{N}(0, \Sigma)$ i.i.d., Y is $m \times n$. Let $S = YY'$. It is easy to check that S has n distinct positive eigenvalues almost surely, and we will assume so from here onward. The density for Y is given by

$$(2\pi)^{-mn/2}(\det \Sigma)^{-n/2} \operatorname{etr}(-\Sigma^{-1}S/2)(dY). \tag{10}$$

Decomposing $Y = H_1DP'$ as in the previous theorem results in $S = H_1LH_1'$, the desired parameterization, where $L \equiv D^2$ and $l_i \equiv L_{ii}$. Note that

$$\bigwedge_{i=1}^n dl_i = 2^n \prod_{i=1}^n D_{ii} \bigwedge_{i=1}^n dD_{ii}$$

and that $\det D = (\det L)^{1/2}$. Replacing (dY) by the right-hand side of (4) and integrating over $(P'\,dP)$ with Corollary 2.1.16 in Muirhead (1982), one therefore obtains the density stated in the theorem. □

The following theorem is a version of Muirhead's Theorem 3.3.1 [Muirhead (1982)] for the case $n = 1$. The general rank-n case is an open problem. [The referee suggested the following general approach. Show that U in Theorem 7 has the same distribution as $(A+B)^{-1/2}A(A+B)^{-1/2}$ and that its moment generating function can be written in terms of a confluent hypergeometric function of matrix argument; see Muirhead (1982). The matrix argument will be a matrix of the same (and thus reduced) rank as A. Reduction formulas in Herz (1955) can then be used to rewrite the moment generating function of U as a moment generating function of lower dimensionality.]

THEOREM 7. *Let $m > 1$ be an integer and let $p > m - 1$. Let A and B be independent, where A is $\mathcal{W}_m(1, \Sigma)$ and B is $\mathcal{W}_m(p, \Sigma)$. Put $A + B = T'T$, where T is an upper-triangular $m \times m$ matrix with positive diagonal elements. Let U be the $m \times m$ symmetric matrix defined by $A = T'UT$. Then $A + B$ and U are* *independent*; $A + B$ *is* $\mathcal{W}_m(p+1,\Sigma)$ *and the density function of* U *on the space* $\mathcal{S}^+_{m,1}$ *with respect to the volume element* (dU) *on this space defined above is*

$$\pi^{(-m+1)/2}\frac{\Gamma_m((p+1)/2)}{\Gamma(1/2)\Gamma_m(p/2)}L^{-m/2}\det(I_m - U)^{(p-m-1)/2}, \tag{11}$$

where $U = H_1LH_1'$, $H_1 \in V_{1,m}$, $L \in \mathbb{R}$.

We conjecture that the density in the general rank-n case is given by

$$\pi^{(-mn+n^2)/2}\frac{\Gamma_m((p+n)/2)}{\Gamma_n(n/2)\Gamma_m(p/2)}\det(L)^{(n-m-1)/2}\det(I_m - U)^{(p-m-1)/2}.$$

PROOF OF THEOREM 7. The proof is almost a verbatim copy of the proof of Muirhead's Theorem 3.3.1 [Muirhead (1982)] and is stated here for reasons of completeness. Find the representation $A = G_1KG_1'$, where $G_1 \in V_{1,m}, K \in \mathbb{R}$. The joint density of A and B is

$$\frac{\pi^{(-m+1)/2}2^{-m(p+1)/2}(\det\Sigma)^{-(p+1)/2}}{\Gamma(1/2)\Gamma_m(p/2)}\operatorname{etr}\left(-\Sigma^{-1}\frac{(A+B)}{2}\right)$$
$$\times K^{-m/2}(\det B)^{(p-m-1)/2}(dA)\wedge(dB).$$

Let $C = A + B$ and note that $(dA) \wedge (dB) = (dA) \wedge (dC)$. Set $C = T'T$, where T is upper triangular with positive diagonal elements, and $A = T'UT$. Find the representation $U = H_1LH_1'$, $H_1 \in V_{1,m}$, $L \in \mathbb{R}$. Theorem 4 implies that

$$(dA)\wedge(dC) = (K/L)^{m/2}(\det T)(dU)\wedge(dC),$$

remembering that T is a function of C alone. Substituting into the density above and collecting terms yields the desired conclusion. □

Analogously to Definition 3.3.2 in Muirhead (1982), we have the following.

DEFINITION 2. A matrix U with density function (11) is said to have the multivariate beta distribution $\mathcal{B}_m(1/2, p/2)$ with parameters $1/2$ and $p/2$.

A matrix V with density function (11) for $U = I_m - V$ is said to have the multivariate beta distribution $\mathcal{B}_m(p/2, 1/2)$ with parameters $p/2$ and $1/2$ (note that one then needs to decompose $I_m - V = H_1LH_1', H_1 \in V_{1,m}, L \in \mathbb{R}$).

It is clear from Theorem 7 that, for $n = 1$, Definition 1 is a special case of Definition 2. Furthermore, by reading Theorem 7 "backwards" and switching the roles of A and B, one obtains another proof for Theorem 1 for the case $n = 1$.

Acknowledgment. I am grateful for useful comments from the referee.

REFERENCES

BOLLERSLEV, T., CHOU, R. Y. and KRONER, K. F. (1992). ARCH modeling in finance—a review of the theory and empirical evidence. *J. Econometrics* **52** 5–59.

FISHER, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* **10** 507–521.

HERZ, C. S. (1955). Bessel functions of matrix argument. *Ann. of Math.* **61** 474–523.

HORN, R. A. and JOHNSON, C. R. (1985). *Matrix Analysis*. Cambridge Univ. Press. (Reprinted in 1987.)

JAMES, A. T. (1954). Normal multivariate analysis and the orthogonal group. *Ann. Math. Statist.* **25** 40–75.

KHATRI, C. G. (1959). On the mutual independence of certain statistics. *Ann. Math. Statist.* **30** 1258–1262.

MUIRHEAD, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

OLKIN, I. and ROY, S. N. (1954). On multivariate distribution theory. *Ann. Math. Statist.* **25** 329–339.

OLKIN, I. and RUBIN, H. (1964). Multivariate beta distributions and independence properties of the Wishart distribution. *Ann. Math. Statist.* **35** 261–269.

SHEPHARD, N. G. (1994). Local scale models—state space alternative to integrated GARCH processes. *J. Econometrics* **60** 181–202.

UHLIG, H. (1992). Bayesian vector autoregressions with stochastic volatility. Unpublished manuscript.

WISHART, J. (1928). The generalized product moment distribution in samples from a normal multivariate population. *Biometrika* **20A** 32–43.

DEPARTMENT OF ECONOMICS
PRINCETON UNIVERSITY
PRINCETON, NEW JERSEY 08544