Research Article

On the Marginal Distribution of the Diagonal Blocks in a Blocked Wishart Random Matrix

Kjetil B. Halvorsen,¹ Victor Ayala,²,³ and Eduardo Fierro⁴

¹Nord University, 8700 Nesna, Norway
²Instituto de Alta Investigación, Universidad de Tarapacá, Antofagasta N. 1520, Arica, Chile
³Universidad Católica del Norte, Región de Antofagasta, Chile
⁴Departamento de Matemáticas, Universidad Católica del Norte, Casilla 1280, Antofagasta, Chile

Correspondence should be addressed to Victor Ayala; vayala@ucn.cl

Received 29 June 2016; Accepted 20 September 2016

Academic Editor: Shamsul Qamar

Copyright © 2016 Kjetil B. Halvorsen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Let \( A \) be a \((m_1 + m_2) \times (m_1 + m_2)\) blocked Wishart random matrix with diagonal blocks of orders \( m_1 \times m_1 \) and \( m_2 \times m_2 \). The goal of the paper is to find the exact marginal distribution of the two diagonal blocks of \( A \). We find an expression for this marginal density involving the matrix-variate generalized hypergeometric function. We became interested in this problem because of an application in spatial interpolation of random fields of positive definite matrices, where this result will be used for parameter estimation, using composite likelihood methods.

1. Introduction

The goal of this paper is to find an exact and useful form for the marginal distribution of the diagonal blocks of a \( 2 \times 2 \) blocked Wishart random matrix. This problem arises in an applied problem, to estimate the parameters of a Wishart random field, which will be reported elsewhere.

Let \( A \) be a \((m_1 + m_2) \times (m_1 + m_2)\) Wishart random matrix, where the diagonal blocks are of orders \( m_1 \times m_1 \) and \( m_2 \times m_2 \), respectively. In our intended application \( m_1, m_2 \) will be small integers (and \( m_1 = m_2 \), but we choose to treat the more general case). Write \( A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \).

Denote the number of freedom parameters by \( n \) and the scale parameter, which is a matrix blocked in the same way as \( A \), by \( \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \). We are mostly interested in the special case \( \Sigma = \begin{pmatrix} \Sigma_0 & \rho \Sigma_a \\ \rho^* \Sigma_a & \Sigma_0 \end{pmatrix} \) where the absolute value of \( \rho \) is less than one, but the general case is not more difficult.

All matrices are real. Notation: we use \( \text{Tr}(A) \) for the trace of the square matrix \( A \) and \( \text{et}(A) = \exp(\text{Tr}(A)) \). We write \( \mathcal{P}(m) \) for the convex cone of real \( m \times m \) positive definite matrices, and we write \( \mathcal{O}(m) \) for the orthogonal group, that is, the set of \( m \times m \) orthogonal matrices. The Stiefel manifold, that is, the set of \( m_1 \times m_2 \) column orthogonal matrices is written as \( \mathcal{V}_{m_1,m_2}(m_2 \leq m_1) \). We indicate the transpose of a matrix by superscript \( \top \).

In the convex cone of positive definite matrices, we use the cone order, defined by \( A < B \) meaning that \( B - A \) is positive definite, written as \( B - A > 0 \). Integrals over cones are written as \( \int_{0}^{I} g(A)(dA) \) meaning the integral is taken over the cone \( 0 < A < I \). The multivariate gamma function is denoted by \( \Gamma_m(a) \) for \( \Re(a) > (m-1)/2 \); see Muirhead [1] for proofs and properties.

In Section 2 we give some background information, especially about the Jacobians which we need to evaluate the integrals. In Section 3 we state our results and give proofs. In Section 4 we give some comments on the result.

2. Background

The single most important reference for background material for this paper is Muirhead [1]. Some results therefrom will not be cited directly.

When doing change of variables in a multiple integral we need to know the Jacobian. Here we will list the ones we need;
most can be found in Muirhead [1] or in Mathai [2]. We are following the notation of Muirhead [1]. First there is a very brief summary.

For any matrix $X$, let $dX$ denote the matrix of differentials $dx_{ij}$. For an arbitrary $m_1 \times m_2$ matrix $X$, the symbol $(dX)$ denotes the exterior product of the $mn$ elements of $dX$:

$$(dX) \equiv \bigwedge_{j=1}^{m_1} \bigwedge_{i=1}^{m_2} dx_{ij}. $$

(1)

If $X$ is a symmetric $m_2 \times m_1$ matrix, the symbol $(dX)$ will denote the exterior product of the $m_2(m_2+1)/2$ distinct elements of $dX$:

$$(dX) \equiv \bigwedge_{1 \leq i < j \leq m_2} dx_{ij}, $$

(2)

with similar definitions for other kinds of structured matrices.

The following invariant form in the orthogonal group represents the Haar measure, $(H^T dH) = \int \mathcal{O}(m_1) dH_i dH_j$. Here $H$ represents an orthogonal matrix. This normalized to have total mass unity is represented by $(dH)$. We also need to integrate over a Stiefel manifold; then $(H^T dH)$ represents a similarly defined invariant form; see Muirhead [1].

Some needed Jacobians are not in Muirhead [1], so we cite those Jacobians here, from Díaz-García et al. [3, 4].

**Lemma 1** (Jacobian of the symmetric square root of a positive definite matrix). Let $S$ and $R$ be in $\mathcal{P}^+(m)$ such that $S = R^2$ and let $\Delta$ be a diagonal matrix with the eigenvalues of $R$ on the diagonal. Then,

$$(dS) = 2^m \det(\Delta) \prod_{i<j}^{m} (\Delta_i + \Delta_j) \, (dR) $$

(3)

$$= \prod_{i < j}^{m} (\Delta_i + \Delta_j) \, (dR). $$

This result can also be found in Mathai [2].

We need the generalized polar decomposition of a rectangular matrix. Let $C$ be $m_1 \times m_2$ rectangular matrix with $m_2 \leq m_1$. Then we always have $C = UH$ where $H$ is positive semidefinite and positive definite if $C$ has full rank, and $U$ is $m_1 \times m_2$ column orthogonal matrix. In that last case, $U$ is unique; see Higham [5].

**Lemma 2** (Generalized polar decomposition). Let $X$ be $m_1 \times m_2$ matrix with $m_1 \geq m_2$ and of rank $m_2$, with $m_2$ distinct singular values. Write $X = UH$, with $U \in \mathcal{O}(m_2)$ and $H \in \mathcal{R}(m_2)$. Then $H$ has $m_2$ distinct eigenvalues. Also let $\Delta$ be the diagonal matrix with the eigenvalues of $H$ on the diagonal. Then

$$(DX) = \det(\Delta)^{m_1-m_2} \prod_{i<j}^{m_2} (\Delta_i + \Delta_j) \, (dH) \, (U^T dU). $$

(4)

Note that since those results are used for integration, the assumption of distinct singular values is unimportant, since the subset where the singular values are equal has measure zero.

### 3. Results

Let us state our main result.

**Theorem 3** (The marginal distribution of the diagonal blocks of a blocked Wishart random matrix with blocks of unequal sizes). Let $A = (A_{12}, A_{12}^T)$ be a $(m_1 + m_2) \times (m_1 + m_2)$ blocked Wishart random matrix, where the diagonal blocks are of sizes $m_1 \times m_1$ and $m_2 \times m_2$, respectively. The Wishart distribution of $A$ has $n \geq m_1 + m_2$ degrees of freedom and positive definite scale matrix $\Sigma = (\Sigma_1, \Sigma_{12}, \Sigma_2)$ blocked in the same way as $A$. The marginal distributions of the two diagonal blocks $A_1$ and $A_2$ have density function given by

$$c \cdot \operatorname{etr}\left\{ -\frac{1}{2} (\Sigma^{-1} A_1 + F^T C_2 F A_1) \right\} \cdot \operatorname{etr}\left\{ -\frac{1}{2} C_2^{-1} A_2 \right\} \cdot \det (A_1)^{(n-m_1-1)/2} \det (A_2)^{(n-m_2-1)/2} $$

(5)

where $C_2 = \Sigma_2 - \Sigma_{12}^T \Sigma_1^{-1} \Sigma_{12}$, $F = C_2^{-1} \Sigma_{12}^T \Sigma_1^{-1}$, and $G = A_1^{1/2} F A_1^{1/2} A_2^{1/2} F A_1^{1/2}$. $c$ is defined and $F$ is the generalized matrix-variate hypergeometric function, as defined in Muirhead [1].

Note that the definition of the matrix-variate hypergeometric function is by a series expansion, which is convergent in all cases we need; see Muirhead [1]. The rest of this section consists of a proof of this theorem.

Introduce the following notation: the Schur complements of $\Sigma = (\Sigma_1, \Sigma_{12}, \Sigma_2)$ is $C_1 = \Sigma_1 - \Sigma_{12}^T \Sigma_{12}^{-1}$ $\Sigma_{12}$ and $C_2 = \Sigma_2 - \Sigma_{12}^T \Sigma_{12}^{-1}$. Then define $F = C_2^{-1} \Sigma_{12}^T \Sigma_1^{-1}$. In the following we will be using some standard results on blocked matrices without quoting them.

The Wishart density function of $A$ written as a function of the blocks is

$$c \cdot \operatorname{etr}\left\{ -\frac{1}{2} (\Sigma^{-1} A_1 + F^T C_2 F A_1 - 2F^T A_{12} + C_2^{-1} A_2) \right\} \cdot \det (A_1)^{\frac{n}{2}} \det (A_2)^{\frac{n}{2}} \det (A_{12})^y $$

(6)

where $c^{-1} = 2^{(m_1 + m_2)} n / (m_1 + m_2 - 1)$ and $y = (n - m_1 - m_2 - 1)/2$. In the following we will work with the density concentrating on the factors depending on $A_{12}$. To prove the theorem we need to integrate out the variable $A_{12}$. The other variables, which are constant under the integration, will be concentrated in one constant factor. So we repeat formula (6) written as a differential form with the constants left out

$$K_1 \cdot \operatorname{etr}(F A_{12}) \det (A_2 - A_{12}^{\top} A_{12}^{-1} A_{12}^T) \, (dA_{12}) $$

(7)

where $K_1 = c \cdot \operatorname{etr}(-[(1/2)(C_2^{-1} A_1 + F^T C_2 F A_1) c^{-1} (-1/2) C_2^{-1} A_2]) \, (dA_1)$. Now, to find the marginal distribution of the diagonal blocks, we need to integrate over the off-diagonal block $A_{12}$. Under this integration the value of the
diagonal blocks $A_1$ and $A_2$ will remain fixed, and the region of integration will be a subset of $\mathbb{R}^{m_1 \times m_2}$ consisting of the matrices $A$ such that the block matrix $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ is positive definite. This seems like a complicated set, but we can give a simple description of it using the polar decomposition of a matrix. Note that this is one of the key observations for the proof, and this author has not seen any use of this observation earlier.

Now we need to assume that $m_1 \geq m_2$. For the opposite inequality a parallel development can be given, using the other factorization $A = \det(A_2) \det(A_1 - A_1^{-1} A_{21} A_{22}^{-1} A_{12})$.

From, for instance, Theorem 1.2.1 in Zhang [6] it follows that the region of integration is the set

$$\left\{ A_{12} \in \mathbb{R}^{m_1 \times m_2} : 0 < A_{12}^T A_{21} A_{22}^{-1} A_{12} < A_2 \right\}. \quad (8)$$

Introduce $E = A_2^{-1/2} A_{12} A_2^{-1/2}$ where we use the usual symmetric square root. Then in terms of the new variable $E$ the region of integration becomes

$$\{ E^T \in \mathbb{R}^{m_1 \times m_2} : 0 < EE^T < I \} \quad (9)$$

and with the generalized polar decomposition in the form $E^T = UP$ with $P \in \mathcal{S}^+$, $U \in \mathcal{V}_{m_1, m_2}$, $EE^T = P^2$ so the region of integration can be written as

$$\{ P \in \mathcal{S}^+(m_2), \ U \in \mathcal{V}_{m_1, m_2}, \ 0 < P^2 < I \} \quad (10)$$

which is a Cartesian product of a cone interval with a Stiefel manifold.

The Jacobian of the transformation from $A_{12}$ to $E$ is

$$(dE) = (dE^T) = \det(A_2)^{-m_1/2} \det(A_1)^{-m_2/2} (dA_{12}).$$

The Jacobian of the polar decomposition $E^T = UP$ is $$(dE^T) = (d\Delta)^{m_1-m_2} \prod_{i=1}^{m_2}(\Delta_i + \Delta) (dP)(U^T dU),$$

where $\Delta$ is a diagonal matrix with the eigenvalues of $P$ on the diagonal; see Lemma 2. A last transformation will be useful. Define $P^2 = X$. The Jacobian of this transformation is $$(dX) = 2^{m_2} \det(\Delta) \prod_{i=1}^{m_2}(\Delta_i + \Delta) (dP); \Delta \text{ as above}. \text{ See Lemma 1.}$$

Applying this transformation the integral of (7) can be written as

$$K_2 \cdot \int_{\mathcal{V}_{m_1, m_2}} \det\left( X^{1/2} A_2^{1/2} FA_1^{1/2} U \right) \det(I - X)^\gamma \cdot \det(X)^{(m_2-m_1-\gamma)/2} (dX)(U^T dU), \quad (11)$$

where the constant

$$K_2 = 2^{-m_2} \det(\Delta)^{1/2} \det(A_2)^{m_2/2} \det(A_1)^{-m_2/2}. \quad (12)$$

We are ready to perform the integration over the Stiefel manifold. For this purpose we need a generalization of Theorem 7.4.1 from Muirhead [1], which we cite here.

Let $X$ be $m \times n$ real matrix with $m \leq n$ and $H = [H_1 : H_2]$ $n \times n$ orthogonal matrix, where $H_1$ is $n \times m$. Then

$$\int_{\mathcal{O}(m)} \det(XH_1)(dH) = o \Gamma \left( n, \frac{1}{4}XX^T \right). \quad (13)$$

But we have an integral over the Stiefel manifold, not the orthogonal group, so we need now to generalize the result (13) to an integral over the Stiefel manifold. What we need is the following. Let $\mathcal{V}_{m_1, m_2}$ be the manifold of $m_1 \times m_2$ column orthogonal matrices with $m_2 \leq m_2$, and let $f$ be a function defined on the Stiefel manifold. We can extend this function to a function defined on $\mathcal{O}(m)$ in the following way. Let $U$ be $m_1 \times m_1$ orthogonal matrix, and write it in block form as $[U_1 : U_2]$ such that $U_1 \in \mathcal{V}_{m_1, m_2}$. How can we characterize the set of $U_2$ which is complementing $U_1$ to form an orthogonal manifold? First, let $U_2$ be a fixed but arbitrary matrix complementing $U_1$. Then clearly any other $m_1 \times (m_1 - m_2)$ column orthogonal matrix with the same column space also works. The common column space is the orthogonal complement of the column space of $U_1$. The set of such matrices can be described as $V \in \mathcal{V}_{m_1 - m_2, m_2}$, $V = U_2Q$ for $Q \in \mathcal{O}(m_1 - m_2)$. For this set we write $\mathcal{V}_{H_1, H_2}$. As a set we can identify this with $\mathcal{O}(m_1 - m_2)$. Specifically, we can identify $U_2$ with the very special column orthogonal matrix $\{\theta_{m_1-m_2}Q\}$, where $Q \in \mathcal{O}(m_1 - m_2)$ which clearly forms a proper submanifold of the Stiefel manifold $\mathcal{V}_{m_1 - m_2, m_2}$. The function $f$ can now be extended to the orthogonal group by defining $f(U) = f([U_1 : U_2]) = f(U_1)$ and for the integral we find that

$$\int_{\mathcal{O}(m)} f(U_1)(U^T dU)$$

$$= \int_{\mathcal{V}_{m_1, m_2}} \int_{\mathcal{V}_{m_1 - m_2, m_2}} f([U_1 : U_2])(U^T dU)$$

$$= \int_{\mathcal{V}_{m_1, m_2}} f(U_1)(U_1^T dU_1) \int_{\mathcal{V}_{m_1 - m_2, m_2}} f(\theta_{m_1 - m_2})(Q^T dQ)$$

$$= \det \left( \frac{\det(\theta_{m_1 - m_2})}{\det(\theta_{m_1 - m_2})} \right)$$

$$= \frac{\det(\theta_{m_1 - m_2})}{\det(\theta_{m_1 - m_2})} \int_{\mathcal{V}_{m_1, m_2}} f(U_1)(U_1^T dU_1), \quad (14)$$

Returning to our integral, the integral over the Stiefel manifold occurring in (11) can now be written as

$$\int_{\mathcal{V}_{m_1, m_2}} \det\left( X^{1/2} A_2^{1/2} FA_1^{1/2} U \right) (U^T dU)$$

$$= \frac{1}{\det(\theta_{m_1 - m_2})} \int_{\mathcal{V}_{m_1, m_2}} \det\left( X^{1/2} A_2^{1/2} FA_1^{1/2} U \right) (U^T dU), \quad (15)$$

where $U_1$ consists of the $m_2$ first columns of $U$
\[
\text{Vol}(\Theta(m_1)) \cdot {}_0F_1\left(\begin{array}{c}
m_2 \\
2 \\
\end{array} \middle| \frac{1}{4} A_1^{1/2} F^\top A_1^{1/2} X A_2^{1/2} F A_1^{1/2} \right),
\]

where we did use (13). Here Vol(\(\Theta(m)\)) = \(2^m \pi^{m^2/2} / \Gamma_m(m/2)\) is the volume of the orthogonal group; see Muirhead [1]. The differential form \((dU)\) denotes Haar measure normalized to total mass unity.

Now write \(G = A_1^{1/2} FA_1 F^\top A_1^{1/2}\); then we can write (11) as

\[
K_2 \cdot \frac{\text{Vol}(\Theta(m_1))}{\text{Vol}(\Theta(m_1 - m_2))} \int_0^1 (\det X)^{(m_1 - m_2) - 1/2} (\det(I - X)) Y \cdot {}_0F_1\left(\begin{array}{c}
m_2 \\
2 \\
\end{array} \middle| \frac{1}{4} GX \right) (dX)
\]

and to evaluate this integral we need Theorem 7.2.10 from Muirhead [1]; we do not state it here.

Using this we find a result we need for the integral of a hypergeometric function, by using the series expansion definition of the hypergeometric function and integrating term by term.

**Theorem 4.** If \(Y\) is a symmetric \(m \times m\) matrix one has that

\[
\int_0^1 (\det X)^{a-(m+1)/2} (\det(I - X))^{b-(m+1)/2} \cdot {}_pF_q\left(\begin{array}{c}
a_1, \ldots, a_p \\
b_1, \ldots, b_q \\
\end{array} \middle| XXY \right) (dX)
\]

\[
= \frac{\Gamma_m(a) \Gamma_m(b)}{\Gamma_m(a + b)} \cdot {}_{p+1}F_{q+1}\left(\begin{array}{c}
a_1, \ldots, a_p, a \\
b_1, \ldots, b_q, a + b \\
\end{array} \middle| Y \right)
\]

so both degrees of the hypergeometric function are raised by one.

The proof is a simple calculation that we leave out.

Now using (18) to calculate (17) we get, finally, the result

\[
K_2 \cdot \frac{\text{Vol}(\Theta(m_1))}{\text{Vol}(\Theta(m_1 - m_2))} \frac{\Gamma_m(m_2/2) \Gamma_m((n - m_1)/2)}{\Gamma_m(n/2)} \cdot {}_1F_2\left(\begin{array}{c}
m_2 \\
n \\
2 \cdot 2 \\
\end{array} \middle| \frac{1}{4} G \right)
\]

but note that one pair of upper and lower arguments to the hypergeometric function are equal with those arguments canceled.

With a little algebra we complete the proof of our main theorem.

4. Some Final Comments

To help interpret our main result, we calculated the conditional distribution of the matrix \(A_1\) given the matrix \(A_2\). We will not give the full details of the calculation here but only give the result. The density of \(A_1\) given that \(A_2 = a_2\) has the density given by

\[
\frac{1}{2^m \pi^{m^2/2} \Gamma_m(n/2) \Gamma_m((m - m_2)/2)} \cdot \det(A_1)^{(m - m_2)/2} \cdot \text{etr}\left(\frac{1}{2} C_1^{-1} A_1 \right)
\]

\[
\cdot {}_0F_1\left(\begin{array}{c}
n \\
2 \\
\end{array} \middle| \frac{1}{4} \Omega C_1^{-1} A_1 \right).
\]

where we have given the conditional density only for the special case \(\Sigma = \left( \begin{smallmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{smallmatrix} \right)\).

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Competing Interests**

Kjetil B. Halvorsen was partially supported by Proyecto Mecesup UCN0711 and Convenio de Cooperación Minera Escondida, Universidad Católica del Norte. Victor Ayala partially was supported by Proyecto Fondecyt no. 1100375 and no. 1150292. Eduardo Fierro was partially supported by Proyecto Mecesup UCN0711.
References

[1] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, 1982.

[2] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*, World Scientific, 1997.

[3] J. A. Diaz-García and G. González-Farías, “Singular random matrix decompositions: Jacobians,” *Journal of Multivariate Analysis*, vol. 93, no. 2, pp. 296–312, 2005.

[4] J. A. Diaz-García, R. Gutierrez Jaimez, and K. V. Mardia, “Wishart and pseudo-Wishart distributions and some applications to shape theory,” *Journal of Multivariate Analysis*, vol. 63, no. 1, pp. 73–87, 1997.

[5] N. J. Higham, *Functions of Matrices, Theory and Computation*, SIAM, 2008.

[6] F. Zhang, Ed., *The Schur Complement and Its Applications*, vol. 4, Springer, New York, NY, USA, 2005.

[7] P. Koev and A. Edelman, “The efficient evaluation of the hypergeometric function of a matrix argument,” *Mathematics of Computation*, vol. 75, no. 254, pp. 833–846, 2006.

[8] R. W. Butler and A. T. Wood, “Laplace approximation for Bessel functions of matrix argument,” *Journal of Computational and Applied Mathematics*, vol. 155, no. 2, pp. 359–382, 2003.
