C. R. Acad. Sci. Paris, Ser. I 354 (2016) 623–627

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Probability theory/Statistics

A weak characterization of real Wishart matrices by quadratic forms

Une caractérisation faible des matrices de Wishart réelles par les formes quadratiques

Gabriel Fraisse [a], Sylvie Viguier-Pla [a,b]

[a] *Université de Perpignan, IUT, Domaine d'Auriac, Carcassonne, France*
[b] *Équipe de statistique et probabilités, IMT, UMR 5219, université Paul-Sabatier, Toulouse, France*

ARTICLE INFO

Article history:
Received 12 May 2011
Accepted after revision 14 March 2016
Available online 14 April 2016

Presented by the Editorial Board

ABSTRACT

Let M be a random symmetric real p-matrix of Wishart distribution with k degrees of freedom and scale parameter Σ. The distribution of M can usually be characterized by the distribution of $({}^t u_1 M u_1, \ldots, {}^t u_p M u_p)$, for any Σ-orthogonal basis $(u_1, \ldots, u_p)$ of $\mathbb{R}^p$. We propose to weaken this characterization, showing that, when $k < p$, it is sufficient to know the distribution of $({}^t u_1 M u_1, \ldots, {}^t u_{k+1} M u_{k+1})$.

RÉSUMÉ

Soit M une p-matrice aléatoire réelle symétrique de loi de Wishart à k degrés de liberté et de paramètre d'échelle Σ. On peut caractériser la loi de M par la loi de $({}^t u_1 M u_1, \ldots, {}^t u_p M u_p)$, pour toute base Σ-orthogonale $(u_1, \ldots, u_p)$ de $\mathbb{R}^p$. Nous proposons une caractérisation plus faible de la loi de M, montrant que, si $k < p$, il suffit de connaître la loi de $({}^t u_1 M u_1, \ldots, {}^t u_{k+1} M u_{k+1})$.

1. Introduction

In multivariate statistics, the quadratic forms issued from Gaussian random variables (r.v.'s) lead implicitly to the Wishart distribution. Representation and characterization of Wishart random matrices is being the object of many works for different uses, as, for example, the computation of moments [3], or time series modeling [6]. A large discussion on the dimension in multivariate models can be found in [1]. This work can be especially useful in a context of large dimension, as the characterization of the Wishart distribution can be weakened in a reduced dimension.

E-mail addresses: fraisse@univ-perp.fr (G. Fraisse), sylvie.viguier-pla@math.univ-toulouse.fr (S. Viguier-Pla).

http://dx.doi.org/10.1016/j.crma.2016.03.011
1631-073X/

We will denote by ${}^t a$ the transpose of any real matrix (or real vector) a. Let p be a strictly positive integer. The identity p-matrix is denoted I_p and the columns of I_p denoted $e_1, \dots, e_p$. The distribution of any random variable U (real, vector or matrix) will be denoted $\mathcal{L}(U)$.

Classically, a Wishart distribution can be defined as follows [2].

Definition 1.1. Let X be a random (k, p)-matrix of k Gaussian rows, centered, identically distributed, independent, of covariance matrix Σ. The distribution of ${}^t X X$ is named Wishart distribution with k degrees of freedom and with scale parameter Σ. This distribution is denoted $\mathcal{W}(k, \Sigma)$.

The second part of this paper will introduce a characterization of the Wishart distribution. The four following paragraphs will expose tools in order to weaken this characterization, that is, respectively, an extension of the Cholesky decomposition, the Haar measure, the Gaussian random matrices, and the Wishart distribution. Finally, the seventh paragraph will conclude with the main announced result.

2. A characterization of Wishart matrices by quadratic forms

When M is a random p-matrix of Wishart distribution $W(k, \Sigma)$, for any vector u such that ${}^t u S u = 1$, the random variable ${}^t u M u$ is chi-square distributed and for any vectors u, v such that ${}^t u S v = 0$, the random variables ${}^t u M u$ and ${}^t v M v$ are independent [5]. We will use these properties to characterize a Wishart distribution.

Theorem 2.1. *Let k and p be strictly positive integers, M be a random symmetric p-matrix and Σ be a non-null symmetric semi-definite positive p-matrix. Then $\mathcal{L}(M) = \mathcal{W}(k, \Sigma)$ if and only if the following properties are verified:*

(a) *if u is a p-vector such that ${}^t u \Sigma u = 0$, then ${}^t u M u = 0$;*
(b) *if u is a p-vector such that ${}^t u \Sigma u \neq 0$, then $\mathcal{L}(\frac{{}^t u M u}{{}^t u \Sigma u}) = \chi^2(k)$ (chi-square distribution with k degrees of freedom);*
(c) *if $(u_1, \dots, u_p)$ is a Σ-orthogonal p-basis, then the r.v.'s ${}^t u_1 M u_1, \dots, {}^t u_p M u_p$ are independent.*

The proof of this theorem needs the two following lemmas.

Lemma 2.1. *Let $p' = rank(\Sigma)$. There exists a non-singular p-matrix A such that, if $p' = p$, ${}^t A \Sigma A = I_p$, and if $p' < p$, ${}^t A \Sigma A = \begin{pmatrix} I_{p'} & 0_{p',p-p'} \\ 0_{p-p',p'} & 0_{p-p',p-p'} \end{pmatrix}$, where $0_{i,j}$ is the null (i, j)-matrix, for any strictly positive integers i and j. Furthermore, $\mathcal{L}(M) = \mathcal{W}(k, \Sigma)$ if and only if $\mathcal{L}({}^t A M A) = \mathcal{W}(k, {}^t A \Sigma A)$.*

Proof. We obtain the columns of A exchanging and normalizing vectors of any Σ-orthogonal p-basis. □

Lemma 2.2. *If properties (a) and (b) of Theorem 2.1 are satisfied and if $\Sigma = \begin{pmatrix} I_{p'} & 0_{p',p-p'} \\ 0_{p-p',p'} & 0_{p-p',p-p'} \end{pmatrix}$, then M is of type $\begin{pmatrix} M' & 0_{p',p-p'} \\ 0_{p-p',p'} & 0_{p-p',p-p'} \end{pmatrix}$, where M' is a random symmetric p'-matrix. Furthermore, $\mathcal{L}(M) = \mathcal{W}(k, \Sigma)$ if and only if $\mathcal{L}(M') = \mathcal{W}(k, I_{p'})$.*

Proof. Suppose that *(a)* and *(b)* are satisfied. Let $M_{i,j}$ be the (i, j)th term of M, where $j \neq i$ and $j > p'$. Property *(a)* implies $M_{j,j} = 0$. From properties *(a)* and *(b)*, we have $\mathcal{L}({}^t(e_i + e_j) M (e_i + e_j)) = \mathcal{L}({}^t(e_i - e_j) M (e_i - e_j)) = \mathcal{L}(M_{i,i})$ and $\mathcal{L}(M_{i,i} + 2M_{i,j}) = \mathcal{L}(M_{i,i} - 2M_{i,j}) = \mathcal{L}(M_{i,i})$. The mathematical expectation of $M_{i,j}$, the covariance of $(M_{i,i}, M_{i,j})$, and the variance of $M_{i,j}$ are null. So $M_{i,j} = 0$. □

Proof of Theorem 2.1. Properties *(a)*, *(b)* and *(c)* are obviously necessary. Suppose now that they are verified. If $\Sigma = I_p$, then they fix the characteristic function f of M. Indeed, let m be a symmetric p-matrix, and $m = r d {}^t r$ its spectral decomposition, where d is a diagonal matrix $\mathrm{diag}(\alpha_1, \dots, \alpha_p)$, and $r = (r_1 | \dots | r_p)$ is an orthogonal p-matrix. Then $f(m) = \mathrm{E}(\exp(\mathrm{trace}(\mathrm{i} m M))) = \Pi_{j=1}^{p} \mathrm{E}(\exp(\mathrm{i} \alpha_j {}^t r_j M r_j))$.

In the general case, from the previous and from Lemmas 2.1 and 2.2, $\mathcal{L}(M) = \mathcal{W}(k, \Sigma)$. □

We will show that, if $p > k$, this set of properties remains a necessary and sufficient condition if we replace the third property by the following:

(c′) for any Σ-orthogonal free family $(u_1, \dots, u_{k+1})$ of p-vectors, the r.v.'s ${}^t u_1 M u_1, \dots, {}^t u_{k+1} M u_{k+1}$ are independent.

3. An extension of the Cholesky factorization

Any definite positive symmetric matrix σ is factorizable in a unique way as $\sigma = {}^t\tau\tau$, where τ is a triangular upper matrix with strictly positive diagonal terms. This is the *Cholesky factorization*. We will denote $\tau = \mathrm{Chol}(\sigma)$.

Let $q \in \{1, \dots, p\}$. For any (p, q)-matrix μ of rank q, the *Schmidt orthonormalization* process applied to the columns of μ leads to the matrix $\mu(\mathrm{Chol}({}^t\mu\mu))^{-1}$, which we denote $\mathrm{Schmidt}(\mu)$. This matrix is such that ${}^t\mathrm{Schmidt}(\mu)\mathrm{Schmidt}(\mu) = I_q$.

We will need the following property.

Property 3.1. *Let a be a (p, k)-matrix and let $q = rank(a)$. There exists a (p, q)-matrix r and a (q, k)-matrix b such that $a = rb$ and ${}^t rr = I_q$.*

Proof. We extract a free q-family $(\mu_1, \dots, \mu_q)$ among the columns of a. Each one of the k columns of a is a linear combination of the μ_j's. So, setting $\mu = (\mu_1|\dots|\mu_q)$, there exists a (q, k)-matrix ν such that $a = \mu\nu$. Then $a = rb$, where $r = \mathrm{Schmidt}(\mu)$ and $b = (\mathrm{Chol}({}^t\mu\mu))\nu$. □

Let $p' = \min(k, p)$ and $\mathcal{S}(k, p)$ be the set of the symmetric p-matrices whose upper left p'-block is definite positive. We can now define on $\mathcal{S}(k, p)$ an application Chol_k such that $\mathrm{Chol}_k = \mathrm{Chol}$ when $k = p$.

Definition 3.1. Let m be a matrix of $\mathcal{S}(k, p)$. We define the k-extended Cholesky factor of m and denote it $\mathrm{Chol}_k(m)$, according to the following way:

if $k < p$, there exists a non-singular k-matrix a, a $(k, p-k)$-matrix b and a $(p-k)$-matrix c such that $m = \begin{pmatrix} {}^t aa & {}^t ab \\ {}^t ba & c \end{pmatrix}$;
then $\mathrm{Chol}_k(m) = {}^t\mathrm{Schmidt}(a)(a|b)$; this definition is intrinsic: ${}^t\mathrm{Schmidt}(a)(a|b) = (\mathrm{Chol}({}^t aa)|{}^t((\mathrm{Chol}({}^t aa))^{-1})\,{}^t ab)$;
if $k = p$, then $\mathrm{Chol}_k(m) = \mathrm{Chol}(m)$;
if $k > p$, then $\mathrm{Chol}_k(m) = \begin{pmatrix} \mathrm{Chol}(m) \\ 0_{k-p,p} \end{pmatrix}$.

Property 3.2. *Let m be a matrix of $\mathcal{S}(k, p)$. Then ${}^t\mathrm{Chol}_k(m)\mathrm{Chol}_k(m) = m$ if and only if there exists a (k, p)-matrix x such that ${}^t xx = m$.*

4. The Haar measure

Let $\mathcal{O}(k)$ be the compact group of the orthogonal k-matrices, and ξ be its Borel σ-field.

The following property is a recall of what can be found in [4].

Property 4.1. *There exists a unique probability measure μ on $\mathcal{O}(k)$, which is left-translation-invariant: $\mu(UE) = \mu(E)$, for any U of $\mathcal{O}(k)$ and E of ξ. It is named the Haar probability measure, and it is also right translation invariant.*

With this property, we can define the Haar-type matrices of $\mathcal{O}(k)$.

Definition 4.1. If R is a random matrix taking values in $\mathcal{O}(k)$ (r.o. k-matrix), then we say that R is of Haar-type (H.r.o.) if its distribution is the Haar probability measure, i.e. if and only if the two propositions, which are equivalent, are satisfied:

i) For any s of $\mathcal{O}(k)$, $\mathcal{L}(sR) = \mathcal{L}(R)$.
ii) For any s of $\mathcal{O}(k)$, $\mathcal{L}(Rs) = \mathcal{L}(R)$.

We will need the following trivial property of the H.r.o. matrices.

Property 4.2. *Let R be a H.r.o. k-matrix. For any r.o. k-matrix S independent of R, RS and SR are H.r.o. k-matrices and are independent of S.*

5. The Gaussian random matrices

Let us first recall the definition of a Gaussian vector and a Gaussian matrix.

Definition 5.1. A random p-vector U is said to be Gaussian if, for any p-vector u, the random variable ${}^t uU$ is Gaussian.

Definition 5.2. A random (k, p)-matrix X is said to be Gaussian if the random kp-vector of its terms is Gaussian.

This lets us deduce the following.

Property 5.1. *A random (k, p)-matrix X is Gaussian if and only if, for any (p, k)-matrix a, the random variable trace(Xa) is Gaussian.*

In particular, we will consider random (k, p)-matrices whose terms are independent and identically distributed as centered and reduced Gaussian r.v.'s. We will denote by $\mathcal{S}N(k, p)$ their distribution.

Property 5.2. *Let r be a matrix of $\mathcal{O}(k)$ and X a random $\mathcal{S}N(k, p)$-matrix. Then $\mathcal{L}(rX) = \mathcal{S}N(k, p)$.*

Proof. The columns X_i of X are independent and isotropic, i.e. such that $\mathcal{L}(rX_i) = \mathcal{L}(X_i)$. Hence, $\mathcal{L}(rX) = \mathcal{L}(X)$. □

Property 5.3. *The rank of any random $\mathcal{S}N(k, p)$-matrix is* $\min(k, p)$.

Proof. The rank of any q-family of almost surely non-null independent isotropic random q-vectors is q. □

6. The Wishart distribution $\mathcal{W}(k, I_p)$

In this part, we suppose that M is a random $\mathcal{W}(k, I_p)$-matrix, according to Definition 1.1.

Property 6.1. *$\mathrm{Chol}_k(M)$ is well defined and ${}^t\mathrm{Chol}_k(M)\mathrm{Chol}_k(M) = M$.*

Proof. From Property 5.3, for any random $\mathcal{S}N(k, p)$-matrix X, $\mathrm{Chol}_k({}^tXX)$ is well defined; moreover, ${}^t\mathrm{Chol}_k({}^tXX)\mathrm{Chol}_k({}^tXX) = {}^tXX$. □

Property 6.2. *If Y is a random (k, p)-matrix such that ${}^tYY = M$, then, for any H.r.o. matrix R independent of Y, $\mathcal{L}(RY) = \mathcal{SN}(k, p)$.*

Proof. We first show this property for $Y = \mathrm{Chol}_k(M)$.

Let suppose that there exists a random $\mathcal{SN}(k, p)$-matrix X such that $M = {}^tXX$.

If $p \geq k$, let X' be the random k-matrix of the k first columns of X. If $p < k$, let X' be the random matrix X completed by a random $\mathcal{SN}(k, k - p)$-matrix independent of X.

In both cases, $\mathcal{L}(X') = \mathcal{SN}(k, k)$. From Property 5.3, $rank(X') = k$. Let T be the r.o. k-matrix Schmidt(X'). Then ${}^tTX = Y$. For any orthogonal k-matrix s, from Property 5.2, $\mathcal{L}(sTY) = \mathcal{L}(sX) = \mathcal{L}(X)$. Since for any orthogonal k-matrix s, $\mathcal{L}(sTY) = \mathcal{SN}(k, p)$, for any r.o. k-matrix S independent of $TY = X$, $\mathcal{L}(STY) = \mathcal{SN}(k, p)$.

Let S be a H.r.o. k-matrix independent of (X, T). Then ST is independent of (X, T) and hence of $M = {}^tXX$, because for any possible value (x, t) of (X, T), $\mathcal{L}(ST/(X, T) = (x, t)) = \mathcal{L}(St/(X, T) = (x, t)) = \mathcal{L}(St) = \mathcal{L}(S)$. We have found a H.r.o. k-matrix $R = ST$ independent of M such that $\mathcal{L}(RY) = \mathcal{SN}(k, p)$.

Now we are able to deal with the general case. Let Y be a random (k, p)-matrix such that ${}^tYY = M$ and R be a H.r.o. k-matrix independent of Y. In the same way as we above obtained from X a r.o. k-matrix T such that ${}^tTX = \mathrm{Chol}_k(M)$, we can obtain from Y a r.o. k-matrix T' such that ${}^tT'Y = \mathrm{Chol}_k(M)$, with (Y, T') independent of R. In the same way as ST was independent of M, RT' is independent of M. Then, from the beginning of the proof, $\mathcal{L}(RY) = \mathcal{L}(RT'\mathrm{Chol}_k(M)) = \mathcal{SN}(k, p)$. □

We have now the tools to weaken the characterization of the Wishart distribution.

7. A weak characterization of Wishart matrices

Let k and p be two integers such that $0 < k < p$, and let M be a random symmetric p-matrix such that:

(i) for any unit p-vector u, $\mathcal{L}({}^tuMu) = \chi^2(k)$;
(ii) for any orthonormal $(k+1)$-family of p-vectors $(u_1, \dots, u_{k+1})$, the r.v.'s ${}^tu_1Mu_1, \dots, {}^tu_{k+1}Mu_{k+1}$ are independent.

Property 7.1. *Let q be an integer such that $1 \leq q \leq k + 1$. For any (p, q)-matrix r such that ${}^trr = I_q$, $\mathcal{L}({}^trMr) = \mathcal{W}(k, I_q)$.*

Proof. The symmetric random q-matrix ${}^trMr$ has the characteristic properties of a Wishart matrix as it has been presented in Theorem 2.1. □

Property 7.2. *$\mathrm{Chol}_k(M)$ is well defined and ${}^t\mathrm{Chol}_k(M)\mathrm{Chol}_k(M) = M$.*

Proof. From Property 7.1, $\mathcal{L}({}^{t}(e_1|\ldots|e_k)M(e_1|\ldots|e_k)) = \mathcal{W}(k, I_k)$. From Property 5.3, $\text{Chol}_k(M)$ is well defined. Let denote $\text{Chol}_k(M)$ by $Y = (Y_1|\ldots|Y_p)$.

For $q \in \{k+1,\ldots,p\}$, let $M' = {}^{t}(e_1|\ldots|e_k|e_q)M(e_1|\ldots|e_k|e_q)$. From Property 7.1, $\mathcal{L}(M') = \mathcal{W}(k, I_{k+1})$. Moreover, $(Y_1|\ldots|Y_k|Y_q) = \text{Chol}_k(M')$; so, from Property 6.1, ${}^{t}Y_qY_q = {}^{t}e_qMe_q$. For $k+1 \le q < q' \le p$, let $f = \frac{1}{\sqrt{2}}(e_q + e_{q'})$ and $F = Yf = \frac{1}{\sqrt{2}}(Y_q + Y_{q'})$. In the same way as ${}^{t}Y_qY_q = {}^{t}e_qMe_q$, ${}^{t}FF = {}^{t}fMf$; so ${}^{t}Y_qY_{q'} = {}^{t}e_qMe_{q'}$. We conclude: ${}^{t}YY = M$. □

Property 7.3. $\mathcal{L}(M) = \mathcal{W}(k, I_p)$.

Proof. Let $Y = \text{Chol}_k(M) = (Y_1|\ldots|Y_p)$, let R be a H.r.o. k-matrix independent of Y and let $X = RY = (X_1|\ldots|X_p)$.

Let a be a non-null (p,k)-matrix and let $q = \text{rank}(a)$. Using Property 3.1, we can write $a = rb$, where ${}^{t}rr = I_q$, and, from Property 7.1, $\mathcal{L}({}^{t}rMr) = \mathcal{W}(k, I_q)$. Properties 6.2 and 5.1 imply respectively $\mathcal{L}(Xr) = \mathcal{L}(RYr) = \mathcal{S}N(k,q)$ and that $\text{trace}(Xrb)$ is Gaussian. Since for any (p,k)-matrix a $\text{trace}(Xa)$ is Gaussian, X is Gaussian.

From Property 7.1, for any (i,j) such that $1 \le i < j \le p$, $\mathcal{L}({}^{t}(e_i|e_j)M(e_i|e_j)) = \mathcal{W}(k, I_2)$. Then, as ${}^{t}(e_i|e_j)M(e_i|e_j) = {}^{t}(Y_i|Y_j)(Y_i|Y_j)$, Property 6.2 implies $\mathcal{L}((X_i|X_j)) = \mathcal{L}(R(Y_i|Y_j)) = \mathcal{SN}(k,2)$. So all the terms of X are centered, reduced, Gaussian r.v.'s of null covariances, hence independent. Then $\mathcal{L}(X) = \mathcal{SN}(k,p)$ and $\mathcal{L}(M) = \mathcal{W}(k, I_p)$. □

From Property 7.3, Lemma 2.1 and Lemma 2.2, we obtain the main result of our work.

Theorem 7.1. *Let k and p be two strictly positive integers, $q = \min(k+1, p)$, M be a random symmetric p-matrix, and Σ be a non-null semi-definite positive symmetric p-matrix. Then $\mathcal{L}(M) = \mathcal{W}(k, \Sigma)$ if and only if the following properties are all satisfied:*

- *if u is a p-vector such that ${}^{t}u\Sigma u = 0$, then ${}^{t}uMu = 0$;*
- *if u is a p-vector such that ${}^{t}u\Sigma u \neq 0$, then $\mathcal{L}(\frac{{}^{t}uMu}{{}^{t}u\Sigma u}) = \chi^2(k)$;*
- *if $(u_1,\ldots,u_q)$ is a Σ-orthogonal free q-family, then the r.v.'s ${}^{t}u_1Mu_1,\ldots,{}^{t}u_qMu_q$ are independent.*

References

[1] T. Bodnar, S. Mazur, K. Podgórski, Singular inverse Wishart distribution and its application to portfolio theory, J. Multivar. Anal. 143 (2016) 314–326.
[2] M.L. Eaton, Multivariate Statistics, Wiley, New York, 1983.
[3] P. Graczyk, G. Letac, H. Massam, The hyperoctahedral group, symmetric group representations and the moments of the real Wishart distribution, J. Theor. Probab. 18 (1) (2005) 1–42.
[4] L. Nachbin, The Haar Integral, D. Van Nostrand Company, Inc., Princeton, NJ, Toronto, London, 1965.
[5] G. Saporta, Probabilités, Analyse des données et statistique, Technip, Paris, 2006.
[6] Y. Xiao, Y.-C. Ku, P. Bloomfield, S.K. Ghosh, On the degrees of freedom in MCMC-based Wishart models for time series data, Stat. Probab. Lett. 98 (2015) 59–64.