Ann. Inst. Statist. Math.
Vol. 47, No. 4, 785–799 (1995)

CRAIG-SAKAMOTO'S THEOREM FOR THE WISHART DISTRIBUTIONS ON SYMMETRIC CONES*

G. LETAC[1] AND H. MASSAM[2]

[1] *Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 31062 Toulouse, France*
[2] *Department of Mathematics and Statistics, York University, North York, Ontario, Canada M3J 1P3*

(Received May 27, 1994; revised January 30, 1995)

Abstract. A version of Craig-Sakamoto's theorem says essentially that if X is a $N(0, I_n)$ Gaussian random variable in $\mathbb{R}^n$, and if A and B are (n, n) symmetric matrices, then $X'AX$ and $X'BX$ (or traces of AXX' and BXX') are independent random variables if and only if $AB = 0$. As observed in 1951, by Ogasawara and Takahashi, this result can be extended to the case where XX' is replaced by a Wishart random variable. Many properties of the ordinary Wishart distributions have recently been extended to the Wishart distributions on the symmetric cone generated by a Euclidean Jordan algebra E. Similarly, we generalize there the version of Craig's theorem given by Ogasawara and Takahashi. We prove that if a and b are in E and if W is Wishart distributed, then Trace $a.W$ and Trace $b.W$ are independent if and only if $a.b = 0$ and $a.(b.x) = b.(a.x)$ for all x in E, where the . indicates Jordan product.

Key words and phrases: Jordan algebra, Wishart distributions, exponential families on convex cones.

1. Introduction

This note has been inspired by two papers on the Craig-Sakamoto's theorem, namely Driscoll and Gundberg (1986) and Ogawa (1993). Both are extremely interesting and thoroughly written papers, the second one completing (and sometimes correcting) the first on many points. Our note has no historical aims and we encourage the reader to have a look at these two papers for a detailed history of the subject. We thank both the editor and the anonymous referees for pointing out some inaccuracies in the references of an earlier version.

We are interested here only on the following simplest version of the Craig-Sakamoto's theorem:

THEOREM 1.1. *Let X be a $N(0, I_n)$ Gaussian random variable in $\mathbb{R}^n$ and let*

* Partially supported by NATO grant 92.13.47.

A and B be symmetric (n, n) matrices. Then

$$U = X'AX \quad \textit{and} \quad V = X'BX$$

are independent random variables if and only if $AB = 0$.

This result is not so easy to prove. It has however an easy extension, which replaces $N(0, I_n)$ by $N(0, \Sigma)$ and $AB = 0$ by $A\Sigma B = 0$, and more difficult one, which replaces $N(0, I_n)$ by $N(\mu, I_n)$ or $N(\mu, \Sigma)$: the story is described in Driscoll and Gundberg (1986). Let us now consider another easy extension of Theorem 1.1, which is already pointed out by Ogasawara and Takahashi (1951). Let us observe first that in Theorem 1.1, $X'AX$ equals Trace AXX', and that $W = XX'$ has the Wishart distribution (denoted $W_n(1, I_n)$, following the notation of Muirhead ((1982), p. 67)). Let us also recall that if we denote by C (resp. int C) the cone of (n, n) symmetric positive real matrices (resp. positive definite), if p belongs to

$$\{1, 2, \ldots, n-1\} \cup (n-1, +\infty) \tag{1.1}$$

and if Σ is in int C, then a random matrix W taking its values in C has Wishart distribution $W_n(p, \Sigma)$ if, for all matrices S such that $\Sigma^{-1} + S$ is in int C, one has

$$\mathbb{E}\left(\exp -\frac{1}{2} \operatorname{Trace} SW\right) = \det(I_n + S\Sigma)^{-p/2}. \tag{1.2}$$

As a generalization of what is done for the gamma distribution, p in (1.1) is called the *shape parameter* and $(2\Sigma)^{-1}$ in int C is the *scale parameter* of $W_n(p, \Sigma)$.

Let us now write the trivial extension of Theorem 1.1 mentioned above:

THEOREM 1.2. *Let p be in* (1.1), *let W be a random variable with distribution $W_n(p, I_n)$ and A and B be symmetric (n, n) matrices. Then*

$$U = \operatorname{Trace}(AW) \quad \textit{and} \quad V = \operatorname{Trace}(BW)$$

are independent if and only if $AB = 0$.

PROOF. Let $c > 0$ be such that $-c < t < c$ and $-c < s < c$ imply that $I_n + tA$, $I_n + sB$, $I_n + tA + sB$ are in Ω. Then it follows from (1.2) that U and V are independent if and only if for $|t|$ and $|s| < c$ one has

$$(\operatorname{Det}(I_n + tA + sB))^{-p/2} = (\operatorname{Det}(I_n + tA))^{-p/2}(\operatorname{Det}(I_n + sB))^{-p/2}. \tag{1.3}$$

Condition (1.3) for p is fulfilled if and only if it is fulfilled for $p = 1$, and from Theorem 1.1, this is equivalent to $AB = 0$. □

During the last ten years, works on the classification of the natural exponential families in $\mathbb{R}^d$ with the simplest variance functions have called attention to extensions of the Wishart distributions to cones other than the cone of positive matrices. Let C be a closed convex cone in a Euclidean space E ($\langle \vec{x}, \vec{y} \rangle$ will denote the scalar product in E) such that the interior int C is not empty and such that $C \cap (-C) = \{\vec{o}\}$.

The dual cone is:

$$C^* = \{\vec{y} \in E; \langle \vec{x}, \vec{y} \rangle \geq 0 \text{ for all } \vec{x} \text{ in } C\}. \tag{1.4}$$

One can show easily that

$$L(\theta) = \int_C \exp(-\langle \vec{\theta}, \vec{x} \rangle) d\vec{x} \tag{1.5}$$

is finite if $\vec{\theta}$ belongs to the interior of C^* (see e.g. Rothaus ((1968), p. 165) for a generalization).

In analogy with (1.1), denote by Λ the set of $\lambda > 0$ such that there exists a positive measure μ_λ on C such that for all $\vec{\theta}$ in int C^* one has

$$(L(\vec{\theta}))^\lambda = \int_C \exp(-\langle \vec{\theta}, \vec{x} \rangle) \mu_\lambda(d\vec{x}). \tag{1.6}$$

Then, for $(\lambda, \vec{\theta})$ in $\Lambda \times$ int C^*, the probability on C

$$W(\lambda, \vec{\theta})(d\vec{x}) = (L(\vec{\theta}))^{-\lambda} \exp(-\langle \vec{\theta}, \vec{x} \rangle) \mu_\lambda(d\vec{x}) \tag{1.7}$$

could be called a *Wishart distribution in the C-sense.*

The distribution (1.7) is a genuine generalization of the ordinary Wishart distribution: if E is the space $H_n(\mathbb{R})$ of symmetric (n, n) matrices, endowed with the scalar product $\langle A, B \rangle = \text{Trace}\, AB$, and if C is $\bar{\Omega}$, the cone of positive matrices, then in fact $C^* = C$, Λ is (1.1) multiplied by the factor $\frac{1}{n+1}$ and, for Σ in int C, the equality:

$$W_n(p, \Sigma) = W\left((n+1)p, \frac{\Sigma^{-1}}{2}\right)$$

links (1.7) with the traditional notation of Wishart distribution on $H_n(\mathbb{R})$.

With this definition (1.7) of the Wishart distribution on a general convex cone of $\mathbb{R}^n$, the problem of the extension of Theorem 1.2 arises. Observe first that $X \mapsto f(X)$ is a linear form on $H_n(\mathbb{R})$ if and only if there exists S in $H_n(\mathbb{R})$ such that $f(X) = \text{Trace}(SX)$. Thus Theorem 1.2 can be seen as a characterization of independent pairs of linear forms of a Wishart distribution in $H_n(\mathbb{R})$. More generally, let W be a random variable in C with distribution defined by (1.7), let a and b be in E and define

$$U = \langle a, W \rangle \quad \text{and} \quad V = \langle b, W \rangle;$$

we may ask for a characterization of the pairs such that U and V are independent. Clearly this is equivalent to

$$L(\theta)L(\theta + ta + sb) = L(\theta + ta)L(\theta + sb) \tag{1.8}$$

for all (t,s) in $\mathbb{R}^2$ such that $\theta+ta$, $\theta+sb$, $\theta+ta+sb$ are in int C^* (here θ is in int C^* and L is defined by (1.5)). Note the independence of condition (1.8) with respect to λ.

The aim of the present note is to solve this problem in the particular case where int C is an irreducible symmetric cone. These cones have already appeared in statistics (see e.g. Jensen (1988)) and their Wishart distributions retain some flavor of the Gaussian origin of the ordinary Wishart. The natural exponential families associated to them have specially nice properties (see Casalis (1990, 1991), Letac (1994), Massam (1994) and Massam and Neher (1994)). We recall a few basic things about them in Sections 2 and 3. Section 4 states and proves our generalization of Craig-Sakamoto's theorem to these Wishart distributions.

2. Symmetric cones and Jordan algebras

One can find several references in the English literature on Jordan algebras. A classical one is Jacobson (1968), but Chapter 1 of Satake (1980) contains a lot of information, references and exercises. Finally the book by Faraut and Koranyi (1994), an elaboration of the notes by Faraut (1988), is an excellent reference for our purposes.

A closed convex cone C in a Euclidean space E is said to be *symmetric* if int C is not empty, $C\cap -C=\{\vec{o}\}$, $C^*=C$ and if the group G of automorphisms of int C acts transitively, i.e. if for all x and y in int C, there exists a linear automorphism g of E such that $g(\text{int } C)=\text{int } C$ and $g(x)=y$. If $E=E_1\oplus E_2$, where E_1 and E_2 are orthogonal with positive dimension, if C_1 and C_2 are symmetric cones of E_1 and E_2, then $C=C_1+C_2$ is also a symmetric cone and C is said to be *irreducible* if there is no such pair (C_1,C_2).

There are only 5 kinds of irreducible symmetric cones. Let us describe them; if K is the algebra $\mathbb{R}$, or $\mathbb{C}$, or $\mathbb{H}$ (Quaternions), or $\mathbb{O}$ (Octonions), denote by $E=H_n(K)$ the space of Hermitian (n,n) matrices with entries in K. An element a in $H_n(K)$ is said to be positive if for all x in K^n, $x'(ax)\geq 0$. $H_{n,+}(K)$ denotes the cone of these positive elements. Up to isomorphism, and with some overlap due to isomorphism, the five kinds of irreducible symmetric cones are

(1) $H_{n,+}(\mathbb{R})$, $\quad n\geq 1$,
(2) $H_{n,+}(\mathbb{C})$, $\quad n\geq 2$,
(3) $H_{n,+}(\mathbb{H})$, $\quad n\geq 2$,
(4) $H_{3,+}(\mathbb{O})$,
(5) the Lorentz cone, i.e.
$\{\vec{x}=(x_1,\ldots,x_n)'\in\mathbb{R}^n; x_1\geq(x_2^2+\cdots+x_n^2)^{1/2}\}, \qquad n\geq 2.$

Actually to each symmetric cone one associates essentially one and only one Euclidean Jordan algebra E (see Faraut (1988), Theorem III.3.1). Let us recall that a Euclidean Jordan algebra is a Euclidean space E with scalar product $\langle a,b\rangle$ and a bilinear map

$$E\times E\to E \qquad (a,b)\mapsto a.b$$

called "Jordan product" with the following properties

(i) the map is symmetric, i.e. $a.b = b.a$,

(ii) there exists e in E such that $a.e = a$,

(iii) $\langle a, b.c\rangle = \langle a.b, c\rangle$ for all a, b, c in E,

(iv) $(a.b).(c.d) + (a.d).(b.c) + (a.c).(b.d) = (a.(c.d)).b + (a.(b.c)).d + (a.(b.d)).c$,
for all a, b, c, d in E.

For a shape E as given above, one can prove (Faraut (1988), Chapter III) that $C = \{a.a; a \in E\}$ is symmetric and that conversely every symmetric cone can be built in that way. When $E = H_n(\mathbb{R})$, the Jordan product is

$$A.B = \frac{1}{2}(AB + BA), \tag{2.1}$$

where AB stands for the ordinary product of matrices. Formula (2.1) holds also in cases (2), (3) and (4). When $E = \mathbb{R}^n$ the Jordan product

$$\vec{x}.\vec{y} = (x_1y_1 + \cdots + x_ny_n, x_1y_2 + y_1x_2, \ldots, x_1y_n + y_1x_n)$$

yields the Lorentz cone of example (5).

Finally a Euclidean Jordan algebra is said to be *simple* if it does not contain a non trivial ideal, i.e. linear subspace I such that $0 < \dim I < \dim E$ and such that the image of $I \times E$ by $(a, b) \mapsto a.b$ is in I. From Faraut ((1988), Chapter III Section 5), $C = \{a.a; a \in E\}$ is irreducible if and only if E is simple.

3. Determinant, trace and Wishart distribution on an irreducible symmetric cone

Let E be a simple Jordan algebra. There are two important polynomial functions, det. and trace, defined on E and with values in $\mathbb{R}$ (see Faraut (1988), Chapter II). For $H_n(\mathbb{R})$ and $H_n(\mathbb{C})$ they coincide with the ordinary determinant and trace of real or complex matrices. See details in Casalis (1990) for $H_n(\mathbb{H})$ and $H_n(\mathbb{O})$. For example 5, $\det(x_1, \ldots, x_n) = x_1^2 - x_2^2 - \cdots - x_n^2$ and $\mathrm{Trace}(x_1, \ldots, x_n) = x_1$.

There are also 3 integers (n, d, r) called the structural constants of E; n is $\dim E$, r is called the rank of E, d is called the Peirce constant and $n = r + \frac{d}{2}r(r-1)$.

For instance, r is the order of matrices in the examples (1), (2), (3), (4) $H_r(K)$, and d is the dimension of K over $\mathbb{R}$, respectively 1, 2, 4, 8. For the Lorentz cone $d = n - 2$, $r = 2$. Needless to say, (d, r) characterises E up to isomorphism.

The trace of $a.b$ is proportional to the scalar product. From now we choose the following normalization of the scalar product: the unit element e, which always satisfies $\mathrm{Trace}\, e = r$, must be such that $\langle e, e\rangle = r$. This implies that $\mathrm{Trace}\, a.b = \langle a, b\rangle$.

Using these quantities, we can write explicitly the Laplace transform (1.5) for an irreducible symmetric cone C.

For all θ in int C, we have:

$$L(\theta) = \int_C \exp(-\langle \theta, x\rangle) d\vec{x} = K_C (\det \theta)^{-n/r} \tag{3.1}$$

where K_C is a constant independent of θ. Thus (1.8) is equivalent to:

$$\det(\theta)\det(\theta + ta + sb) = \det(\theta + ta)\det(\theta + sb) \tag{3.2}$$

for all (t, s) in $\mathbb{R}^2$. Since det is a polynomial, equality (3.2) holds in $\mathbb{R}^2$ if it holds in a neighborhood of $(0, 0)$.

Let us also recall that the set Λ described in (1.6) is, in terms of d and r:

$$\Lambda = \left\{\frac{dr}{2n}, \frac{2dr}{2n}, \frac{3dr}{2n}, \dots, (r-1)\frac{dr}{2n}\right\} \cup \left((r-1)\frac{dr}{2n}, +\infty\right), \tag{3.3}$$

i.e. int $C \to \mathbb{R}$ $\theta \mapsto (\det\theta)^{-pd/2}$ is the Laplace transform of some positive measure on C if and only if p belongs to $\{1, 2, \dots, r-1\} \cup (r-1, +\infty)$.

This result is due to Gindikin (1975). It had been conjectured a long time before by Lévy (1948) for the cone $H_2(\mathbb{R})$: Shanbhag (1988) gives a clever proof of it; it can be easily generalized to symmetric cones (see Casalis and Letac (1994) for this generalization and some bibliographical comments).

4. Craig-Sakamoto's theorem on an irreducible symmetric cone

THEOREM 4.1. *Let E be a simple Euclidean Jordan algebra and C be an irreducible symmetric cone. Let a and b be in E, let λ be in the set* (3.3) *and let X be a random variable with Wishart distribution $W(\lambda, e)$ on C. Consider*

$$U = \mathrm{Trace}(a.X) \quad \textit{and} \quad V = \mathrm{Trace}(b.X).$$

The following four statements are equivalent

(i) *U and V are independent,*

(ii) *for all (t, s) in $\mathbb{R}^2$ $\det(e + ta + sb) = \det(e + ta)\det(e + sb)$,*

(iii) *$a.b = 0$ and for all x in E $a.(b.x) - b.(a.x) = 0$,*

(iv) *there exists an idempotent c in E such that $a \in V(c, 1)$ and $b \in V(c, 0)$.*

Comments. With (3.1) and (3.2), equivalence between (i) and (ii) is almost trivial. Condition (iii) is the closest to the condition $AB = 0$ in Theorem 1.1. But note that in $H_r(\mathbb{R})$, the condition $A.B = 0$, i.e. $AB + BA = 0$ does not imply $AB = 0$, whereas $AB + BA = 0$ plus $A.(B.X) - B.(A.X) = 0$ for all X (i.e. $(AB - BA)X = X(AB - BA)$ for all X in $H_r(\mathbb{R})$) is equivalent to $AB = 0$. Condition (iv) in the case of $H_r(\mathbb{R})$ alludes to the fact that in a Euclidean space, the product of two symmetric endomorphisms a and b is 0 if and only if there exists an orthonormal basis $f_1, \dots, f_r$ such that

$$[a]_f^f = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad [b]_f^f = \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix}.$$

Equivalence between (iii) and (iv) is easy, as well as the implication (iv) $\Rightarrow$ (ii). Finally, like in Theorem 1.1, the more delicate point is (ii) $\Rightarrow$ (iv). There are essentially three different proofs of the necessary condition in Theorem 1.1, the one given by Ogawa (1949), the one given by Matusita (1949)—it was rediscovered independently by Lancaster (1954)—and finally the one given by Ogasawara and Takahashi (1951). These two proofs given in 1949 are historically the first correct proofs of Theorem 1.1. Olkin ((1990), p. 247), also contains variations of the method of Ogasawara and Takahashi. We shall give two proofs of (ii) $\Rightarrow$ (iv), imitating first Ogasawara and Takahashi, then Matusita-Lancaster. Actually Ogawa proves a slightly stronger result, which says in essence that if A and B are 2 symmetric (r, r) real matrices then

$$\det(I_n + t(A+B)) = \det(I_n + tA)\det(I_n + tB)$$

for all t in $\mathbb{R}$ if and only if $AB = 0$: it lacks a clear probabilistic interpretation.

PROOF OF THEOREM 4.1.

$(i) \Rightarrow (ii)$

Denote for simplification $p = \lambda d(r+1)/2$. Then for t and s small enough, from (3.1)

$\mathbb{E}(\exp -tU) = (\det(e + ta))^{-p}$

$\mathbb{E}(\exp -sV) = (\det(e + sb))^{-p}$

$\mathbb{E}(\exp -tU - sV) = (\det(e + ta + sb))^{-p}$.

This identity (ii) holds for t and s small. Since both sides are polynomials in (t, s), it holds for (t, s) in $\mathbb{R}^2$.

$(ii) \Rightarrow (i)$

Similar.

For the remainder of the other equivalence proofs we adopt the following notation. Given a in E, there exists a sequence $(c_1, \dots, c_k)$ of orthogonal primitive idempotents, with $k \leq r$, such that $a = \alpha_1 c_1 + \cdots + \alpha_k c_k$, where the real numbers $\alpha_1 \cdots \alpha_k$ are all different from 0. Let c be the idempotent equal to $c_1 + \cdots + c_k$. Finally if $V(c, \lambda)$ for $\lambda = 0,\ 1,\ 1/2$ denotes the eigenspace of the endomorphism of E $x \mapsto c.x$, we denote by $x = x_0 + x_{1/2} + x_1$ the Peirce decomposition of a given x in E, with x_λ in $V(c, \lambda)$; the functions "determinant" defined on the two subalgebras $V(c, 1)$ and $V(c, 0)$ will be denoted by $\det_1$ and $\det_0$ respectively.

$(iii) \Rightarrow (iv)$

If we apply (iii) to $x = c$ we obtain

$$O = a.(b.c) - b.(a.c) = a.\left(b_1 + \frac{1}{2}b_{1/2}\right) - (b_0 + b_{1/2} + b_1).a.$$

Thus $0 = \frac{1}{2}a.b_{1/2}$. Since $b_{1/2}$ is in $V(c, 1/2)$, we can write

$$b_{1/2} = \sum_{1 \leq i \leq k < j \leq r} b_{ij} \quad \text{with } b_{ij} \quad \text{in } V\left(c_i, \frac{1}{2}\right) \cap V\left(c_j, \frac{1}{2}\right).$$

Thus

$$0 = a.b_{1/2} = \sum_{e,i\leq k<j} \alpha_e c_e.b_{ij} = \sum_{i\leq k<j} \alpha_i c_i.b_{ij} = \frac{1}{2}\sum_{i\leq k<j} \alpha_i b_{ij}.$$

Since the b_{ij} are independent this implies $\alpha_i b_{ij} = 0$, and $b_{ij} = 0$ since $\alpha_i \neq 0$. Thus $b_{1/2} = 0$.

The other part of the hypothesis is $a.b = 0$. This yields $0 = a.(b_1 + b_0) = a.b_1$, and one shows similarly that $b_1 = 0$.

$(iv) \Rightarrow (iii)$

Here b_1 and $b_{1/2}$ are 0, and there exists a sequence $(c_{k+1}, \dots, c_r)$ of orthogonal primitive idempotents such that: $b = \beta_1 c_{k+1} + \cdots + \beta_{r-k} c_r$, where $\beta_1, \dots, \beta_{r-k}$ are in $\mathbb{R}$, and such that $c_{k+1} + \cdots + c_r = e - c$. Thus $a.b = 0$. To see that $a.(b.x) - b.(a.x)$ is zero, we compute this quantity in the three cases: x is $V(c, \lambda)$, with $\lambda = 0$, 1 and 1/2. From different relations between the $V(c, \lambda)$ (see Faraut ((1988), p. 46)) we have $a.(b.x) - b.(a.x) = 0$ when x is in $V(c, \lambda)$ for $\lambda = 1$ or 0. If x is in $V(c, 1/2)$ we write

$$x = \sum_{1\leq i\leq k<j\leq r} x_{ij}, \qquad \text{with } x_{ij} \quad \text{in } V(c_i, 1/2) \cap V(c_j, 1/2).$$

Thus

$$a.x = \sum_{i\leq k<j} \alpha_i x_{ij}, \qquad b.x = \sum_{i\leq k<j} \beta_j x_{ij},$$

from which it follows easily that $a.b.(x) - b.(a.x) = 0$ for all x in E.

$(iv) \Rightarrow (ii)$

We still write $b = \sum_{j=k+1}^{r} \beta_{j-k} c_j$. Thus

$$e + ta + sb = \sum_{i=1}^{k}(1 + t\alpha_i)c_i + \sum_{j=k+1}^{r} (1 + s\beta_{j-k})c_j,$$

$$\begin{aligned}\det(e + ta + sb) &= \prod_{i=1}^{k}(1 + t\alpha_i) \prod_{j=k+1}^{r} (1 + s\beta_{j-k}) \\ &= \det(e + ta)\det(e + sb).\end{aligned}$$

$(ii) \Rightarrow (iii)$ *or* (iv)

Now comes the interesting part of the proof.

1ST PROOF. Ogasawara-Takahashi's method.

We will first recall this method in the real symmetric matrix case: for A and B (r, r) symmetric matrices, and t and s small enough, we write

$$\begin{aligned}\exp - \sum_{n=1}^{\infty} \frac{1}{n}\operatorname{trace}(tA + sB)^n &= \det(I_r - tA - sB) \\ &= \det(I_n - tA)\det(I_n - sB) \\ &= \exp - \sum_{n=1}^{\infty} \frac{1}{n}(t^n \operatorname{Trace} A^n + s^n \operatorname{Trace} B^n).\end{aligned}$$

Thus $\mathrm{Trace}[(tA+sB)^n - t^nA^n - s^nB^n] = 0$ for all $n \geq 1$.

Watching the coefficient of t^2s^2 for $n=4$, one gets

$$2\,\mathrm{Trace}\,AB^2A + \mathrm{Trace}(AB+BA)^2 = 0. \tag{4.1}$$

Since AB^2A and $AB+BA$ are symmetric matrices, this implies $AB=0$.

We now imitate this clever trick for a Jordan algebra E. As usual, we adopt the notation, for x in E and n in $\mathbb{N}$: $x^0 = e$ and $x^{n+1} = x.(x^n)$. We have, for t small enough:

$$\begin{aligned}\det(e-ta) = \prod_{i=1}^{k}(1-t\alpha_i) &= \exp - \sum_{n=1}^{\infty}\frac{t^n}{n}\sum_{i=1}^{k}\alpha_i^n \\ &= \exp - \sum_{n=1}\frac{t^n}{n}\,\mathrm{trace}(a^n).\end{aligned}$$

$$\mathrm{Trace}((ta+sb)^n - t^na^n - s^nb^n) = 0 \qquad \text{for all} \quad n \geq 1. \tag{4.2}$$

We now take $n=4$, look at the coefficient of t^2s^2 in (4.2) and take $t=1$ without loss. One has to compute $(a+sb)^4$ according to the rules of Jordan algebras:

$$\begin{aligned}&(a+sb)^2 = a^2 + 2sa.b + s^2b^2,\\ &(a+sb)^3 = a^3 + s(2a.(a.b) + a^2.b) + s^2(2b.(a.b) + a.b^2) + s^3b^3.\end{aligned}$$

The coefficient of t^2s^2 in (4.2) for $n=4$ is

$$\mathrm{trace}(a.(a.b^2) + 2a.(b.(a.b)) + 2b.(a.b) + b.(b.a^2)) = 0. \tag{4.3}$$

We now have to transform (4.3) into something close to (4.1). To do so, we use (for the first time) the quadratic map P. Let us observe that in general

$$\mathrm{trace}\,P(x)(y) = \mathrm{trace}\,x^2.y \qquad \text{for all} \quad x \text{ and } y \text{ in } E. \tag{4.4}$$

This comes from the fact that

$$\mathrm{trace}((x.y).z) = \mathrm{trace}\,x.(y.z) \qquad \text{for all} \quad x, y \text{ and } z \text{ in } E \tag{4.5}$$

(see Faraut ((1988), pp. 32–33)), and from the very definition of $P(x)(y) = 2x(x.y) - x^2.y$.

Applying (4.5) we get that

$$\mathrm{trace}\,a^2.b^2 = \mathrm{trace}\,a.(a.b^2) = \mathrm{trace}\,b.(b.a^2). \tag{4.6}$$

Applying (4.4) to $x=a$ and $y=b^2$ we get that

$$\mathrm{trace}\,P(a)(b^2) = \mathrm{trace}\,a^2.b^2. \tag{4.7}$$

(4.5) gives also

$$\mathrm{trace}(a.b)^2 = \mathrm{trace}\,a.(b.(a.b)) = \mathrm{trace}\,b.(a.(a.b)). \tag{4.8}$$

Gathering (4.6), (4.7) and (4.8), we finally get the desired consequence of (4.3) as an analogue of (4.1):

$$2\,\mathrm{trace}\, P(a)(b^2) + 4\,\mathrm{trace}(a.b)^2 = 0. \tag{4.9}$$

Since $(a.b)^2$ and b^2 are in the symmetric cone C, since C is preserved by the quadratic maps $P(x)$ and since traces of elements y of C are non negative (and 0 if and only if $y = 0$), we get from (4.9) that $a.b = 0$ and that $P(a)(b^2) = 0$. Finally, we write

$$b^2 = b_1^2 + b_0^2 + b_{1/2}^2 + 2(b_1 + b_0).b_{1/2}. \tag{4.10}$$

Using the general rules

$$\begin{aligned} &P(V(c,1))(V(c,1)) \subset V(c,1),\\ &P(V(c,1))(V(c,0) + V(c,1/2)) = 0, \end{aligned} \tag{4.11}$$

we obtain

$$0 = P(a)(b^2) = P(a)(b_1^2) + P(a)(b_{1/2}^2).$$

Again, since b_1^2 and $b_{1/2}^2$ are in C, we get $P(a)(b_1^2) = 0$.

Since a^{-1} exists in $V(c,1)$ and since $P(a^{-1}) = (P(a))^{-1}$ in $V(c,1)$, we get $b_1^2 = 0$ and $b_1 = 0$. Since $a.b = 0$, we get $a.b_{1/2} = 0$. Writing $b_{1/2} = \sum_{i \le k < j} b_{ij}$ with b_{ij} in $V(c_i, 1/2) \cap V(c_j, 1/2)$, we get $0 = a.b_{1/2} = \sum_{i \le k < j} \alpha_i b_{ij}$, and finally $b_{ij} = 0$ since $\alpha_i \neq 0$ and the b_{ij} are independent. Thus $b_{1/2} = 0$, $a.(b.x) - b.(a.x) = 0$ for all x in E and (iii) is proved.

2ND PROOF. Matusita-Lancaster's method.

We also recall the principle of this method in the case of real symmetric matrices to prove the necessary condition in Theorem 1.1. Without loss of generality, we assume that A and B are written by blocks as follows:

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} B_1 & B_{1/2} \\ B'_{1/2} & B_0 \end{bmatrix}$$

with $A_1 = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_k)$ and $\det A_1 \neq 0$. Thus

$$\frac{1}{t^k} \det(I_n + tA + sB) = \det \begin{bmatrix} \frac{1}{t}(I_k + sB_1) + A_1 & sB_{1/2} \\ \frac{s}{t} B'_{1/2} & I_{r-k} + sB_0 \end{bmatrix}. \tag{4.12}$$

Taking the limit as $t \to +\infty$ in (4.12) and using (ii), we get

$$\alpha_1 \cdots \alpha_k \det(I_{n-k} + sB_0) = \alpha_1 \cdots \alpha_k \det(I_r + sB). \tag{4.13}$$

Comparing the eigenvalues of B_0 and B from (4.13) we get

$$0 = \mathrm{trace}\, B^2 - \mathrm{trace}\, B_0^2 = \mathrm{trace}(B_1^2 + 2B_{1/2}B'_{1/2}).$$

Since B_1^2 and $B_{1/2}B'_{1/2}$ are positive symmetric (k,k) matrices, we obtain $B_1 = 0$, $B_{1/2} = 0$ and $AB = 0$.

To imitate Matusita-Lancaster and equality (4.12), we would need the exact analogue triangular matrices in the context of symmetric cones and this is not available. We use instead the following lemma (a detailed proof is in Massam and Neher ((1994), Proposition 3.3.1)).

LEMMA 4.1. *Let c be an idempotent in the simple Euclidean Jordan algebra E, let $x_0 + x_1 + x_{1/2}$ be the Peirce decomposition of x with respect to c and assume that x^{-1} exists in $V(c,1)$. Then $P(x_{1/2})(x_1^{-1})$ is in $V(c,0)$ and*

$$\det x = \det_1 x_1 \det_0 (x_0 - P(x_{1/2})(x_1^{-1})). \tag{4.14}$$

We apply the lemma to $x_1 = c + ta + sb_1$, $x_0 = e - c + sb_0$ and $x_{1/2} = sb_{1/2}$: for fixed s and for t big enough x_1^{-1} exist; thus from (4.14):

$$\frac{1}{t^k}\det(e + ta + sb) = \det_1\left(\frac{1}{t}(c + sb_1) + a\right)\det_0(x_0 - P(x_{1/2})x_1^{-1}). \tag{4.15}$$

Since, clearly, $x_1^{-1} \to_{t\to+\infty} 0$, taking limits in (4.15) when $t \to +\infty$ and using (ii) we get the analogue of (4.13):

$$\det_1 a \det_0(e - c + sb_0) = \det_1 a \det(e + sb). \tag{4.16}$$

Let $(c_{k+1}, \dots, c_r)$ and $(c'_1, c'_2, \dots, c'_r)$ be two sequences of orthogonal primitive idempotents such that

$$b_0 = \sum_{j=k+1}^{r} \beta_j c_j \quad \text{and} \quad b = \sum_{i=1}^{r} \beta'_i c'_i.$$

From (4.16), since $\det_1 a \neq 0$ we get for all s

$$\prod_{j=k+1}^{r}(1 + \beta_j s) = \prod_{i=1}^{r}(1 + \beta'_i s),$$

$$\operatorname{trace} b_0^2 = \sum_{j=k+1}^{r} \beta_j^2 = \sum_{i=1}^{r}(\beta'_i)^2 = \operatorname{trace} b^2. \tag{4.17}$$

$V(c,1/2)$ is orthogonal to $V(c,0)$ and $V(c,1)$ (thus $\operatorname{trace} b_{1/2}(b_1 + b_0) = 0$). Using this and (4.10), we get from (4.17) that $\operatorname{trace}(b_1^2 + b_{1/2}^2) = 0$. Since b_1^2 and $b_{1/2}^2$ are in C, this implies b_1 and $b_{1/2} = 0$ and (iv) is true.

5. Another result of independence

The following result is easily obtained as a consequence of Theorem 1.1:

THEOREM 5.1. *Let $(\xi, \eta)'$ be a Gaussian centered random variable of $\mathbb{R}^{n+k}$ with covariance matrix:*

$$\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{1/2} \\ \Sigma'_{1/2} & \Sigma_0 \end{bmatrix}.$$

Let A_1 and B_0 be symmetric (n,n) and (k,k) matrices. Assume that Σ, A_1, B_0 are invertible. Then, if $\xi' A_1 \xi$ and $\eta' B_0 \eta$ are independent, one has $\Sigma_{1/2} = 0$.

PROOF. We write $A = \sqrt{\Sigma}\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}\sqrt{\Sigma}$ and $B = \sqrt{\Sigma}\begin{bmatrix} 0 & 0 \\ 0 & B_0 \end{bmatrix}\sqrt{\Sigma}$, and note that

$$X = (\sqrt{\Sigma})^{-1}\begin{pmatrix} \xi \\ \eta \end{pmatrix}$$

has distribution $N(0, I_{n+k})$. An application of Theorem 1.1 shows that $\xi' A_1 \xi$ and $\eta' B_0 \eta$ are independent if and only if $A\Sigma B = 0$. Since $A\Sigma B = \begin{bmatrix} 0 & A_1\Sigma_{1/2}B_0 \\ 0 & 0 \end{bmatrix}$ and since A_1 and B_0 are invertible, the result is proved. □

Just as we passed from Theorem 1.1 to Theorem 1.2 by introducing the ordinary Wishart distribution, we could give an easy extension of Theorem 5.1 with Wishart distributions again. We opt not to do so, but rather to give a generalization in the framework of symmetric cones.

THEOREM 5.2. *Let W be a random variable taking its values in a simple Euclidean Jordan algebra E, Wishart distributed with scale parameter y in int C (C is the symmetric cone of E) and shape parameter p in (3.3), i.e. for θ in int C*

$$\mathbb{E}(\exp - \operatorname{Trace}\theta W) = ((\det y)/(\det(y+\theta)))^{pd/r}.$$

Let c be an idempotent and let a and b be in $V(c,1)$ and $V(c,0)$ respectively. We assume that a and b are invertible with respect to these subalgebras. Then trace$(a.W)$ *and* trace$(b.W)$ *are independent implies that $y_{1/2} = 0$, where $y_{1/2}$ is the projection of y on $V(c, 1/2)$.*

PROOF. Clearly trace$(a.W)$ and trace$(b.W)$ are independent if and only if for all (t, s) in $\mathbb{R}^2$ one has

$$\det y \det(y + sa + sb) = \det(y + sa)\det(y + tb). \tag{5.1}$$

Since y is in int C, and if $y = y_1 + y_0 + y_{1/2}$ is the Peirce decomposition of y, clearly y_1^{-1} and y_0^{-1} exist. Denote $z_0 = P(y_{1/2})y_1^{-1}$.

From Lemma 4.2 it is in $V(c,0)$, and $y_0 - z_0$ is in C. Let us apply Lemma 4.2 to each of the 4 determinants of 5.1. After trite simplifications, we get

$$\begin{aligned}(5.2)\qquad \det_0(y_0 - z_0)\det_0(y_0 - tb - P(y_{1/2})(y_1 - sa)^{-1}) \\ = \det_0(y_0 - z_0 - tb)\det_0(y_0 - P(y_{1/2})(y - sa)^{-1}).\end{aligned}$$

Letting $s \to +\infty$, (5.2) gives

$$(5.3)\qquad \det_0(y_0 - z_0)\det_0(y_0 - tb) = \det y_0 \det(y_0 - z_0 - tb).$$

Since $\det(P(y)(x)) = (\det y)^2(\det x)$ for any x and y in E and $y_0 - z_0$, being in C, has a square root in C, (5.3) becomes

$$(5.4)\qquad \det_0(y_0 - z_0)\det(y_0 - tb) = \det_0 y_0 \det_0(e - tP((\sqrt{y_0 - z_0})^{-1}(s))).$$

For $t = 0$ in (5.4), we get $\det_0(y_0 - z_0) = \det_0 y_0$, or $\det_0(e - P((\sqrt{y_0})^{-1})(z_0))) = 1$. Since $z_0' = P((\sqrt{y_0})^{-1})(z_0)$ and $e - z_0'$ are both in C, the eigenvalues of $e - z_0'$ are in $[0,1]$, their product is 1, therefore $z_0' = z_0 = 0$. Thus $P(y_{1/2})y_1^{-1} = 0$. Writing

$$y_1^{-1} = \sum_{i=1}^{k} \alpha_i c_i \qquad \text{with} \quad \alpha_1 > 0 \;\; \alpha_2 > 0 \cdots \alpha_k > 0,$$

and $c = c_1 + \cdots + c_k$ with $(c_1, \ldots, c_k)$ a sequence of primitive orthogonal idempotents, we have

$$0 = \operatorname{Trace} P(y_{1/2})(y_1^{-1}) = \langle y_{1/2}, y_{1/2} \cdot y_1^{-1} \rangle,$$

by using (4.4) and (4.5). Writing

$$y_{1/2} = \sum_{i \leq k < j} a_{ij},$$

where a_{ij} is in $V(c_i, 1/2) \cap V(c_j, 1/2)$, we get

$$y_{1/2} \cdot y_1^{-1} = 1/2 \sum_{i \leq k < j} \alpha_i a_{ij}$$

and

$$0 = \frac{1}{2} \sum_{i \leq k < j} \alpha_i \|a_{ij}\|^2,$$

since the a_{ij} are orthogonal. Since $\alpha_i > 0$, it follows that $a_{ij} = 0$ and $b_{1/2} = 0$. □

6. Comments

The first author of this present note has been asked by P. Doisy (Toulouse) which extension of Theorem 1.1 to infinite dimensional spaces are possible. The most natural one is to consider symmetric measurable functions a and $b : [0,1]^2 \to \mathbb{R}$ such that a and b are in $L^2([0,1]^2)$, and to consider double integrals with respect to standard Brownian motion

$$U = \int_0^1 \int_0^1 a(x,y)dB(x)dB(y), \qquad V = \int_0^1 \int_0^1 b(x,y)dB(x)dB(y).$$

It has been proved by Ustünel and Zakaï (1989) that U and V are independent if and only if $\int_0^1 a(x,z)b(z,y)dy = 0$ in the $L^2([0,1]^2)$ sense. It also has been generalized to higher multiple integrals in the same paper. Sufficient condition (which was trivial for Theorem 1.1) has got a simpler proof with Kallenberg (1991); we are indebted to Marc Yor for these two references. Needless to say, these infinite dimensional results are using the stochastic calculus of Ito and Malliavin and are no longer elementary.

REFERENCES

Casalis, M. (1990). Familles exponentielles naturelles invariantes par un groupe, Ph.D. Thesis, Laboratoire de Statistique et Probabilités, Université Paul Sabatier, Toulouse.

Casalis, M. (1991). Les familles exponentielles à variance quadratique homogène sont des lois de Wishart sur un cône symétrique, *C.R. Acad. Sc. Paris*, **312**, 537–540.

Casalis, M. and Letac, G. (1994). Characterization of the Jorgensen set in the generalized linear model, *Test*, **3**, 145–161.

Driscoll, M. F. and Gundberg, W. R. (1986). A history of the development of Craig's theorem, *Amer. Statist.*, **40**, 65–70.

Faraut, J. (1988). Algèbres de Jordan et cônes symétriques, Notes d'un Cours d'Eté du Centre International de Mathématiques Pures et Appliquées, Département de Mathématiques, Université de Poitiers.

Faraut, J. and Koranyi, A. (1994). *Analysis on Symmetric Cones*, Oxford University Press, Oxford.

Gindikin, S. (1975). Invariant generalized functions in homogeneous spaces, *J. Funct. Anal. Appl.*, **9**, 50–52.

Jacobson, N. (1968). Structure and representations of Jordan algebras, *Amer. Math. Soc. Collog. Publ.*, **XXXXIX**, Providence, Rhode Island.

Jensen, S. (1988). Covariance hypotheses which are linear in both the covariance and in the inverse covariance, *Ann. Statist.*, **16**, 302–322.

Kallenberg, O. (1991). On an independence criterion for multiple Wiener integrals, *Ann. Probab.*, **19**, 83–85.

Lancaster, H. O. (1954). Traces and cumulants of quadratic forms in the normal variables, *J. Roy. Statist. Soc. Ser. B*, **16**, 247–254.

Letac, G. (1994). Les familles exponentielles statistiques invariantes par les groupes du cône et du paraboloïde de révolution, *Volume in Honor of L. Takacs* (eds. J. Galambos and J. Gani), Applied Probability Trust, Sheffield.

Lévy, P. (1948). The arithmetical character of the Wishart distribution, *Proc. Cambridge Philos. Soc.*, **44**, 255–297.

Massam, H. (1994). An exact decomposition theorem and a unified view of some related distributions for a class of exponential transformation models on symmetric cones, *Ann. Statist.*, **22**, 369–394.

Massam, H. and Neher, E. (1994). On transformations and determinants of Wishart variables on symmetric cones (submitted).

Matusita, K. (1949). Note on the independence of certain statistics, *Ann. Inst. Statist. Math.*, **1**, 79–82.

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

Ogasawara, T. and Takahashi, M. (1951). Independence of quadratic quantities in a normal system, *Journal of Science of the Hiroshima University, Series A*, **15**, 1–9.

Ogawa, J. (1949). On the independence of bilinear and quadratic forms of a random sample from a normal population, *Ann. Inst. Statist. Math.*, **1**, 83–108.

Ogawa, J. (1993). A history of the development of Craig-Sakamoto's theorem viewed from Japanese standpoint, *Proc. Inst. Statist. Math.*, **21**, 47–59 (in Japanese).

Olkin, I. (1990). Interface between statistics and linear algebra, *Proc. Sympos. Appl. Math.*, **40**, 233–256.

Rothaus, O. (1968). Some properties of Laplace transforms of measures, *Trans. Amer. Math. Soc.*, **131**, 163–169.

Sakamoto, H. (1944). On the independence of statistics, Research Memo., **1**, The Institute of Statistical Mathematics, Tokyo (in Japanese).

Satake, I. (1980). *Algebraic Structures on Symmetric Domains*, Iwanami-Shoten, Tokyo and Princeton University Press, Princeton.

Shanbhag, D. N. (1988). The Davidson-Kendall problem and related results on the structure of the Wishart distribution, *Austral. J. Statist.*, **30A**, 272–280.

Ustünel, A. S. and Zakaï, M. (1989). On independence and conditioning on Wiener space, *Ann. Probab.*, **17**, 1441-1453.