The Annals of Statistics
1998, Vol. 26, No. 2, 573–595

QUADRATIC AND INVERSE REGRESSIONS FOR WISHART DISTRIBUTIONS[1]

BY GÉRARD LETAC AND HÉLÈNE MASSAM

Université Paul Sabatier and York University

If U and V are independent random variables which are gamma distributed with the same scale parameter, then there exist a and b in $\mathbb{R}$ such that

$$\mathbb{E}(U \mid U+V) = a(U+V)$$

and

$$\mathbb{E}(U^2 \mid U+V) = b(U+V)^2.$$

This, in fact, is characteristic of gamma distributions. Our paper extends this property to the Wishart distributions in a suitable way, by replacing the real number U^2 by a pair of quadratic functions of the symmetric matrix U. This leads to a new characterization of the Wishart distributions, and to a shorter proof of the 1962 characterization given by Olkin and Rubin. Similarly, if $\mathbb{E}(U^{-1})$ exists, there exists c in $\mathbb{R}$ such that

$$\mathbb{E}(U^{-1} \mid U+V) = c(U+V)^{-1}.$$

Wesołowski has proved that this also is characteristic of gamma distributions. We extend it to the Wishart distributions. Finally, things are explained in the modern framework of symmetric cones and simple Euclidean Jordan algebras.

1. Introduction. Let p and σ be two positive numbers. We denote by

$$\gamma_{p,\sigma}(du) = \exp(-\sigma^{-1}u)\sigma^{-p}u^{p-1}1_{(0,+\infty)}(u)(\Gamma(p))^{-1}\,du \tag{1.1}$$

the gamma distribution on $(0, +\infty)$ with scale parameter σ and shape parameter p. Let U and V be two independent, positive, non-Dirac random variables, and consider the following five properties:

1. $U/(U+V)$ and $U+V$ are independent.
2. There exist two real numbers a and b such that

$$\mathbb{E}(U \mid U+V) = a(U+V), \tag{1.2}$$

$$\mathbb{E}(U^2 \mid U+V) = b(U+V)^2. \tag{1.3}$$

2′. There exist two real numbers a and c such that

$$\begin{aligned}\mathbb{E}(U \mid U+V) &= a(U+V),\\ \mathbb{E}(U^{-1} \mid U+V) &= c(U+V)^{-1}.\end{aligned} \tag{1.4}$$

Received August 1996; revised July 1997.
[1]Supported by NATO Grant 921347 and NSERC Grant A8947.
AMS 1991 *subject classifications*. Primary 62H05; secondary 60E10.
Key words and phrases. Natural exponential families, Wishart distributions, Jordan algebras, conditional moments.

3. There exist p, q and σ such that the distributions of U and V are, respectively $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$.
3′. Same as (3), plus $p > 1$.

It is known that (1), (2) and (3) are equivalent, and that (2′) and (3′) are equivalent, too.

Let us first concentrate on the first group of equivalences; if either (2) or (3) holds, the following formulas are true:

$$a = \frac{p}{p+q}, \qquad b = \frac{(p+1)p}{(p+q+1)(p+q)}. \tag{1.5}$$

The implication (3) ⇒ (1) is the easiest to prove, since if $-p < t < q$ and if $\theta > \sigma^{-1}$, one has

$$\mathbb{E}\Big((UV^{-1})^t \exp(-\theta(U+V))\Big) = \frac{\Gamma(p+t)}{\Gamma(p)} \frac{\Gamma(q-t)}{\Gamma(q)} (1+\theta\sigma)^{-p-q}. \tag{1.6}$$

The proof of (1) ⇒ (3) has been given by Lukacs (1955). However, it has been noted by several authors [e.g., Wang (1981)] that (1) ⇒ (2) is quite simple and that, this together with (2) ⇒ (3), is a little bit easier than the direct proof of (1) ⇒ (3). At the beginning of Section 4, we will give a proof of (2) ⇒ (3) using the variance function of the natural exponential family $\{\gamma_{p,\sigma};\ \sigma > 0\}$.

The most natural generalization of (1.1) is the Wishart distribution on the space E of (r, r) symmetric matrices. Throughout this paper, we will use the same notation as Casalis and Letac (1996). We denote by $\overline{E_+}$ and E_+ the cones of positive and positive definite matrices, respectively, and we write

$$\Lambda = \{1/2, 1, 3/2, \dots, (r-1)/2\} \cup ((r-1)/2, +\infty). \tag{1.7}$$

For p in Λ, there exists a positive measure μ_p on $\overline{E_+}$ such that, if σ belongs to E_+, then the measure on $\overline{E_+}$:

$$\gamma_{p,\sigma}(du) = \exp\big(-\operatorname{trace}(\sigma^{-1}u)\big)(\det\sigma)^{-p}\mu_p(du) \tag{1.8}$$

is a probability distribution. Clearly, if we replace σ by $(\theta + \sigma^{-1})^{-1}$ in (1.8), we get that, for $(\theta + \sigma^{-1})$ in E_+,

$$\int \exp(-\operatorname{trace}(\theta u))\gamma_{p,\sigma}(du) = \det(I_r + \theta\sigma)^{-p}. \tag{1.9}$$

We call $\gamma_{p,\sigma}$ the Wishart distribution with scale parameter p and shape parameter σ. It should be pointed out that, in order to avoid tedious constants, we differ slightly from the traditional notation $W_r(m, \Sigma)$ for Wishart distributions, as found, for instance, in Muirhead (1982). The correspondence between the two notations is simply given by $\gamma_{m/2, 2\Sigma} = W_r(m, \Sigma)$.

A natural question is now the following. Does there exist a generalization, to the Wishart distributions, of the equivalence between (1), (2) and (3)? For the equivalence between (1) and (3), the answer is yes and has been provided by Olkin and Rubin (1962). This paper is difficult to read; Casalis and Letac (1996), using many of Olkin and Rubin's ideas, provides a simpler and clearer proof of (1) ⇒ (3) [or, more accurately, of the corresponding statements for the general Wishart distribution given by (1.8); see Section 5]. The statement (3) ⇒ (1) for the general Wishart distribution is not as easy as (1.6) and is also proved in Casalis and Letac (1996).

The first problem considered in the present paper is the equivalence between (2) and (3) for the Wishart distributions (1.7). Surprisingly enough, as in the one-dimensional case, (2) ⇒ (3) will be much simpler than (1) ⇒ (3). Since, as we will see in Proposition 5.1, the implication (1) ⇒ (2) is also easy to prove, we shall end up with a new and shorter proof of Olkin and Rubin's theorem in Section 5.

Now, finding the correct generalization of (1) to the Wishart distribution was one of the numerous bright ideas of Olkin and Rubin. For (2), the generalization of (1.2) is clear, and we just have to write a trivial extension of a result by Rao (1948) to obtain it. However, the proper generalization of (1.3) is not that obvious, but can be found in the toolbox of the proof of (1) ⇒ (3) given in Casalis and Letac (1996). Replacing the real number U^2 by the square of the symmetric matrix U is not the right idea [see the remarks following Corollary 2.3 for the explicit expression of $\mathbb{E}(U^2 \mid U + V)$]. What we use is a pair of quadratic functions of U valued in the linear space $L_s(E)$ of symmetric endomorphisms of the Euclidean space E of symmetric real (r, r) matrices. The Euclidean structure on E is given by the scalar product $\langle A, B \rangle = \text{trace } AB$. Thus $\dim E = r(r+1)/2$ and $\dim L_s(E) = r(r+1)(r^2+r+2)/8$. These two quadratic functions are: $U \otimes U$, defined by

$$(1.10) \qquad E \to E, \qquad H \mapsto (U \otimes U)(H) = U \,\text{trace}(UH),$$

and $\mathbb{P}(U)$, defined by

$$(1.11) \qquad E \to E, \qquad H \mapsto \mathbb{P}(U)(H) = UHU.$$

For example, if $r = 2$, and if we take

$$U = \begin{bmatrix} a+b & c \\ c & a-b \end{bmatrix},$$

then U belongs to E_+ if and only if $a > (b^2 + c^2)^{1/2}$. In this case, as an orthogonal basis of E we can take the three matrices:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Thus, if $H = xI + yJ + zK$, we have $(U \otimes U)(H) = 2(ax + by + cz)U$ and

$$\mathbb{P}(U)(H) = \begin{bmatrix} (a+b)^2(x+y) + c^2(x-y) + 2(a+b)cz & 2acx + 2bcy + (a^2 - b^2 + c^2)z \\ 2acx + 2bcy + (a^2 - b^2 + c^2)z & (a-b)^2(x+y) + c^2(x+y) + 2(a-b)cz \end{bmatrix};$$

that is, the matrices representative of $U \otimes U$ and $\mathbb{P}(U)$ in the basis $\{I, J, K\}$ are, respectively,

$$2\begin{bmatrix} a^2 & ab & ac \\ ba & b^2 & bc \\ ca & cb & c^2 \end{bmatrix}$$

and

$$\begin{bmatrix} a^2+b^2+c^2 & 2ab & 2ac \\ 2ab & a^2+b^2-c^2 & 2bc \\ 2ac & 2bc & a^2-b^2+c^2 \end{bmatrix}.$$

Let us show that it is natural to replace (1.3) for real symmetric (r,r) matrices by the following: There exists a $(2,2)$ real matrix $b = (b_{ij})_{i,j=1,2}$ such that

$$\begin{aligned} \mathbb{E}(U \otimes U \mid U+V) &= b_{12}(U+V) \otimes (U+V) + b_{12}\mathbb{P}(U+V), \\ \mathbb{E}(\mathbb{P}(U) \mid U+V) &= b_{21}(U+V) \otimes (U+V) + b_{22}\mathbb{P}(U+V). \end{aligned} \tag{1.12}$$

The main idea leading to the consideration of the pair (1.10)–(1.11) can be found in Casalis and Letac (1996), where quadratic maps $Q\colon E \to L_s(E)$ with a certain invariance property were considered. This property can be presented here as follows: For any A in the linear group $GL(r,\mathbb{R})$ of invertible real (r,r) matrices, and for any (U,H) in E^2, one has

$$Q(AU{}^tA)(H) = AQ(U)({}^tAHA)^{t}A \tag{1.13}$$

[a reformulation of (6.1) in Casalis and Letac (1996) for symmetric matrices]. It is easily seen that the functions Q defined by (1.10) or (1.11) fulfill condition (1.13); we shall see in Proposition 2.1 that the linear space of possible Q has dimension 2, and that (1.10)–(1.11) form a basis of this space. Therefore it is natural to consider this pair of quadratic functions of U.

Of course, to justify (1.12), we also have to prove that these two identities are satisfied when U and V are two independent and Wishart-distributed random variables with the same scale parameter σ; that is, we have to prove the implication (3) ⇒ (2). In other words, we have to find the values of a, b_{11}, b_{12}, b_{21} and b_{22} corresponding to a and b in (1.4). We shall actually prove in Section 3 that if p and q are the shape parameters of U and V, and if

$$A(p) = \begin{bmatrix} p^2 & p \\ p/2 & p/2+p^2 \end{bmatrix}, \tag{1.14}$$

then the matrix b in (1.12) is $A(p)(A(p+q))^{-1}$. Such a formula for Wishart distributions has an ancestor in Casalis (1992), Theorem 1. The result in (1.12) can be presented differently if we diagonalize the matrix b. One obtains a pair of identities. Let $p > 1/2$ be an element of Λ, let

$$X(p) = \frac{1}{p(p-1/2)}(U \otimes U - \mathbb{P}(U)),$$
$$Y(p) = \frac{1}{p(p+1)}\left(\frac{1}{2}U \otimes U + \mathbb{P}(U)\right)$$

and let $X(p+q)$ and $Y(p+q)$ be defined in an analogous way by replacing (U, p) by $(U + V, p + q)$. The identities have a martingale flavor and read

$$\mathbb{E}(X(p) \mid U + V) = X(p+q),$$
$$\mathbb{E}(Y(p) \mid U + V) = Y(p+q).$$

Note that, for $p = 1/2$, U is concentrated on the cone of matrices of rank 1 and thus

$$U \otimes U = \mathbb{P}(U).$$

Let us now consider (2′) and (3′). Their equivalence has been observed rather recently in Wesołowski (1990), where

$$c = \frac{p+q-1}{p-1}. \tag{1.15}$$

Our second aim in the present paper is to generalize the equivalence between (2′) and (3′) to the Wishart distributions. Here, things are a little bit simpler than in the quadratic case (2), and the real number U^{-1} can safely be replaced by the inverse matrix U^{-1}. However, the analogue of (2′) has to be supplemented by the consideration of $U^{-1} \otimes U$, the element of the linear space $L(E)$ of endomorphisms of the Euclidean space E of symmetric real (r, r) matrices defined by

$$E \to E, \qquad H \mapsto (U^{-1} \otimes U)(H) = U^{-1}\operatorname{trace}(UH),$$

which necessarily occurs when one considers the differential of $\theta \mapsto \mathbb{E}(U^{-1} \exp\langle \theta, U \rangle)$. For example, if $r = 2$ and if

$$U = \begin{bmatrix} a+b & c \\ c & a-b \end{bmatrix},$$

as above, then the representative matrix of $U^{-1} \otimes U$ in the basis $\{I, J, K\}$ will be

$$\frac{2}{a^2 - b^2 - c^2}\begin{bmatrix} a^2 & ab & ac \\ -ba & -b^2 & bc \\ -ca & -cb & -c^2 \end{bmatrix}.$$

Thus, for symmetric real (r, r) matrix, (1.4) will be replaced by the following: There exist three real numbers c, c_1 and c_2 such that

$$\mathbb{E}(U^{-1} \mid U + V) = c(U+V)^{-1}, \tag{1.16}$$

$$\mathbb{E}(U^{-1} \otimes U \mid U + V) = c_1 id_E + c_2 (U+V)^{-1} \otimes (U+V). \tag{1.17}$$

Statement (3′) is, of course, replaced by a statement with Wishart distributions and $p > (r+1)/2$.

To justify (1.16) and (1.17), we have to prove that these identities are actually satisfied for some c, c_1, c_2, when U and V are independent and have Wishart distributions with the same scale parameter σ [i.e., we have to find the analogue of (1.15)]. We shall prove in Section 6 that if p and q are the shape parameters of U and V, then the constants in (1.16) and (1.17) are

$$c = \frac{p+q-(r+1)/2}{p-(r+1)/2}, \qquad c_1 = \frac{-q}{(p-(r+1)/2)(p+q)},$$
$$c_2 = \frac{p(p+q-(r+1)/2)}{(p-(r+1)/2)(p+q)}.$$

This leads to the rather unexpected consequence, that if U and V are Wishart distributed with the same scale parameter σ, and with shape parameters $p > (r+1)/2$ and q in Λ as defined in (1.7), then the conditional covariance of U^{-1} and U, given $U+V$, is

$$\frac{-q}{(p-(r+1)/2)(p+q)}\, id_E$$

(see Corollary 2.6).

Let us now list the contents of the paper. In Section 2, we state the results which were roughly sketched above. We think that no new results on Wishart distributions should nowadays be presented outside the framework of the Wishart distributions on symmetric cones and simple Euclidean Jordan algebras. Following this policy, as was done, for instance, in Casalis and Letac (1996) or in Massam (1994), we give our results on symmetric cones. It still remains, of course, that our results are new for the classical Wishart distributions. The rank and the Peirce constant of the simple Euclidean Jordan algebra are denoted, respectively, by r and d. The reader who is interested only in symmetric real matrices of order r should set $d = 1$ in the formulas and follow the instructions given in Casalis and Letac (1996), Section 3. Section 2 also contains a proof of the extension of Rao's paper, as well as a proof of the fact that quadratic maps $E \to L_s(E)$ with the invariance property (1.13) form a linear space with dimension 2. Section 3 is computational; we essentially prove the implication (3) ⇒ (2). In Section 4, we give the proof of (2) ⇒ (3) and, as in Casalis and Letac (1996), the main tool is the variance function of the natural exponential family of the Wishart distributions with shape parameter p. In Section 5, we prove the implication (1) ⇒ (2) for Wishart distributions. Section 6 is parallel to Section 3; it is computational and we essentially prove (3′) ⇒ (2′). In Section 7, we prove (2′) ⇒ (3′); the basic tool in that proof is the derivation of Schwartz's distributions. In Section 8, we comment on an unsolved problem raised by a referee.

2. The characterization of Wishart distributions by quadratic and inverse regressions. As explained before, we keep the notation of Casalis and Letac (1996). Let E be a simple Euclidean Jordan algebra, with product $E \times E \to E(a, b) \mapsto a \cdot b$, and scalar product $\langle a, b \rangle = \text{trace}(a \cdot b)$. We denote by $L(E)$ and $L_s(E)$, respectively, the spaces of endomorphisms and symmetric endomorphisms of the Euclidean space E. The identity element of E is e. The closed cone of squares $\{a \cdot a;\ a \in E\}$ is denoted by $\overline{E_+}$, and E_+ is its interior; r is the rank of E, d is its Peirce constant and $n = r + dr(r-1)/2$ is the dimension of E. To avoid trivialities, we always assume $r > 1$. We denote by G the connected component, containing the identity, of the group of linear automorphisms of the linear space E which preserve E_+; we write K for the intersection of G with the orthogonal group of E. The determinant in E is written det: $E \mapsto \mathbb{R}$. We write

$$\Lambda = \{d/2, d, 3d/2, \dots, (r-1)d/2\} \cup ((r-1)d/2, +\infty). \tag{2.1}$$

It is known [Casalis and Letac (1996), Section 3] that if p belongs to Λ, then there exists a positive measure μ_p on $\overline{E_+}$ such that for all θ in E_+ one has

$$\int \exp(-\langle \theta, u \rangle)\, \mu_p(du) = (\det \theta)^{-p}. \tag{2.2}$$

For σ in E_+ and p in Λ, $\gamma_{p,\sigma}$, as defined in (1.8), is called the Wishart distribution with shape parameter p and scale parameter σ. Definitions for objects in Jordan algebras are taken from Faraut and Koranyi (1994), with the exception of the very definition of the Wishart distribution. What they call Wishart distributions are particular cases of our $\gamma_{p,\sigma}$ distributions and are obtained as images of Gaussian distributions on a Euclidean space F by representations of F into E. Thus they choose to generalize χ^2 distributions only, not gamma distributions as we do. For instance, this implies that, in their restricted sense, there exists no Wishart distribution on the exceptional Jordan algebra of dimension 27 corresponding to $r = 3$ and $d = 8$, while they do exist in our sense.

We now introduce two quadratic maps from E to the linear space $L_s(E)$ of symmetric endomorphisms of E defined as follows:

1. For x in E, $x \otimes x$ in $L_s(E)$ is defined by

$$h \mapsto (x \otimes x)(h) = x\langle x, h \rangle, \qquad E \to E.$$

2. For x in E, $\mathbb{P}(x)$ in $L_s(E)$ is defined by

$$h \mapsto \mathbb{P}(x)(h) = 2x \cdot (x \cdot h) - (x \cdot x) \cdot h, \qquad E \to E. \tag{2.3}$$

The map $x \mapsto \mathbb{P}(x)$, $E \to L_s(E)$ is called the quadratic representation. It is known [see Casalis and Letac (1996)] that if p belongs to Λ, then $F_p = \{\gamma_{p,\sigma};\ \sigma \in E_+\}$ is the natural exponential family of Wishart distributions with scale parameter p and variance function

$$V_F(m) = \frac{1}{p}\,\mathbb{P}(m), \tag{2.4}$$

when m varies in the domain of the means $M_{F_p} = E_+$.

To measure the importance for quadratic regression of the pair (2.3), let us prove the following statement (which will be needed later in the proofs of Proposition 3.1 for the first part and Theorem 2.4 for the second).

PROPOSITION 2.1. *Let $\mathcal{Q}$ be the linear space of functions Q: $E_+ \to L(E)$ which are G-invariant, that is, such that for all g in G one has*

$$Q(g(x)) = gQ(x)g^* \tag{2.5}$$

and such that $Q(e)$ is in $L_s(E)$. Then $\mathcal{Q}$ has dimension 2, and the pair $x \mapsto x \otimes x$ and $\mathbb{P}$ defines a basis of $\mathcal{Q}$. Furthermore, if $x \in E$ is such that there exists a in $\mathbb{R}$ with $\mathbb{P}(x) = ax \otimes x$, then either $x = 0$ or $a = 1$ and x is the multiple of some primitive idempotent.

PROOF. It has been observed in Casalis and Letac (1996) after (6.1) that these two polynomials $x \mapsto x \otimes x$ and $\mathbb{P}(x)$ belong to $\mathcal{Q}$. To see that they are linearly independent, we apply them to e and get $x \mapsto (x \otimes x)(e) = x\langle x, e\rangle$ and $\mathbb{P}(x)(e) = x \cdot x$. The results are clearly independent. To see that $\mathcal{Q}$ has dimension 2, consider the linear map

$$\phi\colon \mathcal{Q} \to L_s(E), \qquad Q \mapsto \phi(Q) = Q(e).$$

Then $Q(e)$ is K-invariant, since

$$Q(e) = Q(k(e)) = kQ(e)k^*.$$

This comes from (2.5) and the definition of K. Thus, from Proposition 6.1 of Casalis and Letac (1996) or from Olkin and Rubin (1962), Lemma 1 (6), there exists (λ, μ) in $\mathbb{R}^2$ such that

$$Q(e) = \lambda id_E + \mu e \otimes e.$$

This implies that the dimension of the image of ϕ is ≤ 2 [since it is included in the plane of $L_s(E)$ generated by id_E and $e \otimes e$]. Finally, ϕ is one-to-one: Let us take Q in the kernel of ϕ. Since G acts transitively on E_+, let x be an arbitrary point of E_+ and g in G such that $g(e) = x$. Thus we have

$$Q(x) = Q(ge) = gQ(e)g^* = g0g^* = 0,$$

that is, Q is 0 on E_+ and ϕ is one-to-one. Thus Q and the image of ϕ have the same dimension 2.

To prove the second part of the proposition, assume that $\mathbb{P}(x) = ax \otimes x$ and that x is not 0. Then there exists a sequence of orthogonal primitive idempotents $(c_1, \dots, c_v)$ such that

$$x = \lambda_1 c_1 + \cdots + \lambda_v c_v,$$

with λ_i in $\mathbb{R} \setminus \{0\}$ for $i = 1, \dots, v$. Since for all h in E one has

$$\mathbb{P}(x)(h) = ax\langle h, x\rangle,$$

let us take $h = c_1$. Thus we get

$$\lambda_1^2 c_1 = \mathbb{P}(x)(c_1) = a(\lambda_1 c_1 + \cdots + \lambda_v c_v)\lambda_1,$$

which implies that $v = 1$ and that $a = 1$. The proposition is proved. □

REMARK. Note that, in the previous proposition, imposing that $Q(e)$ be symmetric implies that $Q(x)$ is also symmetric for all x in E_+. Relaxing this condition leads to a surprise: The dimension of $\mathscr{Q}$ is still 2 except in the case $(r, d) = (2, 1)$, that is, in the case of $(2, 2)$ real matrices, for which the dimension of $\mathscr{Q}$ is 3 [see Letac and Massam (1997), where the above result is proved by Neher in an appendix].

We now state the main results of the paper, Theorem 2.2 (the quadratic regression property), Theorem 2.4 (the characterization by quadratic regression), Theorem 2.5 (the inverse regression property) and Theorem 2.7 (the characterization by inverse regression):

THEOREM 2.2. *With the above notation, let σ be in E_+, let p and q be in Λ and let U and V be independent random variables with Wishart distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$, respectively. Then if*

$$A(p) = \begin{bmatrix} p^2 & p \\ pd/2 & p(1-d/2) + p^2 \end{bmatrix} \tag{2.6}$$

and

$$b = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = A(p)(A(p+q))^{-1}, \tag{2.7}$$

one has

$$\begin{aligned} \mathbb{E}(U \otimes U \mid U+V) &= b_{11}(U+V) \otimes (U+V) + b_{12}\mathbb{P}(U+V), \\ \mathbb{E}(\mathbb{P}(U) \mid U+V) &= b_{21}(U+V) \otimes (U+V) + b_{22}\mathbb{P}(U+V). \end{aligned} \tag{2.8}$$

COROLLARY 2.3. *Under the hypothesis of Theorem 2.2, define X and Y in $\mathscr{Q}$ by*

$$X(u) = u \otimes u - \mathbb{P}(u)$$

and

$$Y(u) = \frac{d}{2} u \otimes u + \mathbb{P}(u).$$

Then

$$\mathbb{E}(X(U) \mid U+V) = \frac{p(p-d/2)}{(p+q)(p+q-d/2)} X(U+V),$$

$$\mathbb{E}(Y(U) \mid U+V) = \frac{p(p+1)}{(p+q)(p+q+1)} Y(U+V).$$

PROOF. It is easily seen that the eigenvalues and the line eigenvectors of the matrix b are given by

$$(1,-1)b = \frac{p(p-d/2)}{(p+q)(p+q-d/2)}(1,-1),$$

$$(d/2,1)b = \frac{p(p+1)}{(p+q)(p+q+1)}(d/2,1).$$

To obtain the two identities in the corollary, we multiply (2.8), written in matrix form as

$$\mathbb{E}\left[\begin{array}{c} U\otimes U \\ \mathbb{P}(U) \end{array}\middle| U+V\right] = b\left[\begin{array}{c} (U+V)\otimes(U+V) \\ \mathbb{P}(U+V) \end{array}\right]$$

by the two eigenvectors. $\square$

REMARKS. $X(u) = 0$ if and only if u is proportional to a primitive idempotent, as we have seen in Proposition 2.1. Note that if $p = d/2$ and if U is $\gamma_{d/2,\sigma}$ distributed, then U is almost surely proportional to some primitive idempotent [see Casalis (1990)], and $X(U) = 0$ almost surely. Note also that under the hypothesis of Theorem 2.2, one has

$$\mathbb{E}(U \operatorname{trace} U \mid U+V) = b_{11}(U+V)\operatorname{trace}(U+V) + b_{12}(U+V)^2,$$

$$\mathbb{E}(U^2 \mid U+V) = b_{21}(U+V)\operatorname{trace}(U+V) + b_{22}(U+V)^2$$

[use $\mathbb{P}(u)(e) = u^2$ and apply (2.8) to e].

The next theorem is the converse of Theorem 2.2; its hypothesis is the generalization to Wishart distributions of statement (2) in Section 1.

THEOREM 2.4. *Let U and V be two independent non-Dirac random variables taking their values in $\overline{E_+}$. Assume that $U+V$ is not concentrated on a half line. Assume also that there exists $a \in \mathbb{R}$ such that*

$$\mathbb{E}(U \mid U+V) = a(U+V) \tag{2.9}$$

and that there exists a $(2,2)$ matrix b such that the two equalities (2.8) holds. Then there exist (p,q) in Λ^2 and σ in E_+ such that U and V have distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$. Furthermore, $a = p/(p+q)$ and $b = A(p)(A(p+q))^{-1}$ as in (2.7).

THEOREM 2.5. *With the above notation, let σ be in E_+, let $p > n/r$ and q be in Λ and let U and V be independent random variables with Wishart distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$, respectively. Then for*

$$\begin{aligned} c &= \frac{p+q-n/r}{p-n/r}, \quad c_1 = \frac{-q}{(p-n/r)(p+q)}, \\ c_2 &= \frac{p(p+q-n/r)}{(p-n/r)(p+q)}, \end{aligned} \tag{2.10}$$

one has

$$\mathbb{E}(U^{-1} \mid U + V) = c(U + V)^{-1},$$
$$\mathbb{E}(U^{-1} \otimes U \mid U + V) = c_1 id_E + c_2(U + V)^{-1} \otimes (U + V). \tag{2.11}$$

COROLLARY 2.6. *Under the hypothesis of Theorem* 2.5 *one has*

$$\mathrm{Cov}((U^{-1}, U) \mid U + V) = \frac{-q}{(p - n/r)(p + q)} id_E.$$

PROOF. By Huyghens' formula, for any random variables X and Y taking their values in a Euclidean space, we have

$$\mathrm{Cov}(X, Y) = \mathbb{E}(X \otimes Y) - \mathbb{E}(X) \otimes \mathbb{E}(Y).$$

Applying this formula for the conditional covariance of $(X, Y) = (U^{-1}, U)$ given $U + V$, using (2.10), (2.11) and

$$\mathbb{E}(U \mid U + V) = \frac{p}{p + q}(U + V),$$

we obtain the desired result. □

The last theorem is the converse of Theorem 2.5; its hypothesis is the generalization to Wishart distributions of statement (2′) of Section 1.

THEOREM 2.7. *Let U and V be two independent non-Dirac random variables taking their values in $\overline{E_+}$. Assume that $U + V$ is not concentrated on a half line and that $\mathbb{E}(U^{-1})$ exists. Assume also that there exist a, c, c_1 and c_2 such that* (2.9) *and the two equalities* (2.11) *hold. Then there exist $p > n/r$, q in Λ and σ in E_+ such that U and V have $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$ distributions. Furthermore, $a = p/(p + q)$ and c, c_1 and c_2 are as in* (2.10).

To complete this section, we make a small digression about (2.9), generalizing Rao's result (1948) as follows:

PROPOSITION 2.8. *Let E be a Euclidean space and let U and V be two independent non-Dirac random variables taking their values in E. Assume that the two sets $\Theta(U)$, defined as the interior of*

$$\{\theta; L_U(\theta) = \mathbb{E}(\exp\langle \theta, U\rangle) < \infty\},$$

and $\Theta(V)$, defined similarly, have a nonempty intersection. Assume also that there exists $a \in \mathbb{R}$ such that

$$\mathbb{E}(U \mid U + V) = a(U + V).$$

Then $0 < a < 1$, $\Theta(U) = \Theta(V)$ *and*

$$L_U(\theta) = (L_V(\theta))^{a/(1-a)}. \tag{2.12}$$

PROOF. For θ in $\Theta(U) \cap \Theta(V)$, we have

$$\begin{aligned} L'_U(\theta) L_V(\theta) &= \mathbb{E}(U \exp\langle \theta, U + V \rangle) \\ &= a\mathbb{E}((U+V)\exp\langle \theta, U+V \rangle) = aL'_{U+V}(\theta). \end{aligned}$$

Thus

$$(2.13) \qquad (1-a)\frac{L'_U(\theta)}{L_U(\theta)} = a\,\frac{L'_v(\theta)}{L_V(\theta)}.$$

The principle of maximal analyticity [Kawata (1972)] shows that $\Theta(U) = \Theta(V)$ from (2.13). Now, if a were equal to 0 or 1, (2.13) would imply that U or V are Dirac. Finally, since the differential of $\theta \mapsto L'_U(\theta)/L_U(\theta)$ must be positive, then $1-a$ and a have the same sign and therefore $0 < a < 1$. And (2.12) follows immediately from (2.13). □

3. Proof of Theorem 2.2. The proof of Theorem 2.2 relies on the following two propositions, which are of interest in their own right.

PROPOSITION 3.1. *Let* Ψ: $L_s(E) \to L_s(E)$ *be the linear map defined by*

$$(3.1) \qquad \Psi(x \otimes x) = \mathbb{P}(x)$$

for all x *in* E. *Then*

$$(3.2) \qquad \Psi(\mathbb{P}(x)) = \frac{d}{2}\, x \otimes x + \left(1 - \frac{d}{2}\right)\mathbb{P}(x).$$

PROPOSITION 3.2. *Let* U *be a random variable taking its values in* E *with Wishart distribution* $\gamma_{p,\sigma}$, *as defined in Section* 2 [*or as given in* (1.8)). *Then*

$$(3.3) \qquad \begin{aligned} \mathbb{E}(U \otimes U) &= p^2\sigma \otimes \sigma + p\mathbb{P}(\sigma), \\ \mathbb{E}(\mathbb{P}(U)) &= \frac{pd}{2}\,\sigma \otimes \sigma + \left(p - \frac{pd}{2} + p^2\right)\mathbb{P}(\sigma). \end{aligned}$$

PROOF OF PROPOSITION 3.1. The existence and uniqueness of Ψ were proved in Lemma 6.3 of Casalis and Letac (1996); the notation that we adopt here replaces $\Psi_{\mathbb{P}}$ by Ψ. We observe that the map

$$x \mapsto Q(x) = \Psi(\mathbb{P}(x)), \qquad E_+ \to L_s(E)$$

is G-invariant in the sense of (2.5). To see this, we take g in G and write

$$(3.4) \quad g\mathbb{P}(x)g^* = \mathbb{P}(g(x)) = \Psi(g(x) \otimes g(x)) = \Psi(g(x \otimes x)g^*).$$

The first equality in (3.4) is a standard property of $\mathbb{P}$, the second is by definition of Ψ and the third is obvious. Now, since the set $\{x \otimes x;\ x \in E\}$ generates the linear space $L_s(E)$, we obtain, from (3.4) by linearity, that

$$(3.5) \qquad g\Psi(f)g^* = \Psi(gfg^*).$$

From (3.5) now applied to $f = \mathbb{P}(x)$, we get

$$g\Psi(\mathbb{P}(x))g^* = \Psi(g\mathbb{P}(x)g^*) = \Psi(\mathbb{P}(g(x))).$$

Thus we have proved that

$$x \mapsto \Psi(\mathbb{P}(x)), \qquad E_+ \to L_s(E)$$

is G-invariant in the sense of (2.5). Proposition 2.1 implies that there exist λ and μ in $\mathbb{R}$ such that, for all x in E_+,

$$\Psi(\mathbb{P}(x)) = \lambda\mathbb{P}(x) + \mu x \otimes x. \tag{3.6}$$

Since both members of (3.6) are quadratic polynomials, this equality (3.6) is also true for all x in E. To compute λ and μ, we use (6.12) in Casalis and Letac (1996) by taking $x = e$ in (3.6) and we get (3.2). □

Proof of Proposition 3.2. Writing $A = \mathbb{E}(U \otimes U)$ and $B = \mathbb{E}(\mathbb{P}(U))$ and using $A(p)$ as defined in (2.6), identities (3.3), to be proved, can be written as

$$\begin{bmatrix} A \\ B \end{bmatrix} = A(p)\begin{bmatrix} \sigma \otimes \sigma \\ \mathbb{P}(\sigma) \end{bmatrix}.$$

From (2.2), we have that, for $-\theta$ in E_+, $L_{\mu_p}(\theta) = (\det(-\theta))^{-p}$ and $k_{\mu_p} = \log L_{\mu_p}$. For any quadratic polynomial Q defined on E, it is a general property of Laplace transforms that

$$\int_{\overline{E_+}} Q(u)\exp\langle \theta, u\rangle \mu_p(du) = \left(Q\left(\frac{d}{d\theta}\right)k_{\mu_p}(\theta) + Q\big(k'_{\mu_p}(\theta)\big)\right)L_{\mu_p}(\theta). \tag{3.7}$$

We apply (3.7) to $Q(u) = u \otimes u$ and note that, in this case, $Q(d/d\theta)j = j''$ for any C^2-function j defined on an open set of E, and that $k'_{\mu_p}(\theta) = p(-\theta)^{-1}$ and $k'_{\mu_p}(\theta) = p\mathbb{P}((-\theta)^{-1})$; we obtain the first identity in (3.3), for A, by taking $\sigma = (-\theta)^{-1}$.

The second identity in (3.3) for B is more delicate and will use Proposition 3.1. Let us apply (3.7) to $Q = \mathbb{P}$. To simplify the notation, we write $k(\theta) = -\log\det(-\theta)$ for $-\theta \in E_+$. We get

$$\begin{aligned} &\int_{\overline{E_+}} \mathbb{P}(u)\exp\langle \theta, u\rangle \mu_p(du) \\ &\quad = \left(p\mathbb{P}\left(\frac{d}{d\theta}\right)(k)(\theta) + p^2\mathbb{P}\big((-\theta)^{-1}\big)\right)(\det(-\theta))^{-p}. \end{aligned} \tag{3.8}$$

To compute $\mathbb{P}(d/d\theta)(k)$ in (3.8), we now use (6.11) in Casalis and Letac (1996), which says that Ψ as defined in Proposition 3.1 satisfies $\mathbb{P}(d/d\theta)(j) = \Psi(j'')$. Since $k''(\theta) = \mathbb{P}((-\theta)^{-1})$, Proposition 3.1 can be applied and, with the help of (3.8) with $\sigma = (-\theta)^{-1}$, yields the second formula of (3.3). □

Comments. Recalling the definition of the variance function as given in (2.4), the first formula in (3.3) is nothing but a multivariate version of Huyghens' formula linking the second moment and the variance. The second formula of (3.3) could also be deduced from Theorem 1 of Casalis (1992), whose proof is not really shorter.

We now prove Theorem 2.2. First, we note that, in order to prove

$$\mathbb{E}(Q(U) \mid U+V) = Q_1(U+V)$$

for given functions Q and Q_1, it is enough to show that, for θ in some nonempty open set,

$$\mathbb{E}(Q(U)\exp\langle\theta, U+V\rangle) = \mathbb{E}(Q_1(U+V)\exp\langle\theta, U+V\rangle). \tag{3.9}$$

We apply this principle to the pair

$$\begin{aligned} Q(u) &= \begin{bmatrix} u\otimes u \\ \mathbb{P}(u) \end{bmatrix}, \\ Q_1(u) &= A(p)(A(p+q))^{-1}\begin{bmatrix} u\otimes u \\ \mathbb{P}(u) \end{bmatrix}. \end{aligned} \tag{3.10}$$

Take $-\theta$ in E_+ and, for simplicity, denote

$$\sigma_\theta = (\sigma^{-1} - \theta)^{-1}. \tag{3.11}$$

Using (3.10), (3.3) can be rewritten as

$$\mathbb{E}(Q(U)) = A(p)\begin{bmatrix} \sigma\otimes\sigma \\ \mathbb{P}(\sigma) \end{bmatrix}, \tag{3.12}$$

changing σ into σ_θ in (3.12) gives

$$\mathbb{E}(Q(U)\exp\langle\theta, U\rangle) = A(p)\begin{bmatrix} \sigma_\theta\otimes\sigma_\theta \\ \mathbb{P}(\sigma_\theta) \end{bmatrix} L_U(\theta),$$

$$\mathbb{E}(Q_1(U+V)\exp\langle\theta, U+V\rangle) = A(p)\begin{bmatrix} \sigma_\theta\otimes\sigma_\theta \\ \mathbb{P}(\sigma_\theta) \end{bmatrix} L_{U+V}(\theta).$$

Formula (3.9) follows immediately and Theorem 2.2 is proved.

4. Proof of Theorem 2.4. We first show the implication (2) ⇒ (3), as given in Section 1. The proof of Theorem 2.4 will be an extension of this simple case.

As usual, we write $L_U(\theta) = \mathbb{E}(e^{\theta U})$ and $k_U = \log L_U$; k'_U exists on $(-\infty, 0)$, since $U \ge 0$. We write $\chi = k_{U+V}$. The generalization of Rao's theorem, given in Proposition 2.8, and (1.2) imply that

$$k_U = a\chi. \tag{4.1}$$

Thus, from (1.3), we derive that, for $\theta < 0$,

$$\begin{aligned} (a\chi''(\theta) + a^2\chi'^2(\theta))L_{U+V}(\theta) &= L''_U(\theta)L_V(\theta) \\ &= \mathbb{E}(U^2\exp(\theta(U+V))) \\ &= b\mathbb{E}((U+V)^2\exp(\theta(U+V))) \\ &= (b\chi''(\theta) + b\chi'^2(\theta))L_{U+V}(\theta). \end{aligned}$$

This gives $(a-b)\chi'' = (b-a^2)\chi'^2$. Since $0 < a < 1$ (see Proposition 2.5), $a - b = b - a^2 = 0$ is impossible. Since $U+V$ is non-Dirac, $a - b = 0$ is impossible. Thus, if $\lambda = (b - a^2)/(a - b)$, one has

$$\chi''(\theta) = (\chi'(\theta))^2/\lambda. \tag{4.2}$$

The equality (4.2) is identical to (2.11) in Casalis and Letac (1996). It says that the variance function of the natural exponential family (NEF) generated by the distribution of $U + V$ is $V(m) = m^2/\lambda$ on its domain of the means. Thus $U + V$, U and V follow the gamma distributions with shape parameters λ, $a\lambda = p$ and $(1 - a)\lambda = q$, respectively; the scale parameter σ is the same for all variables.

Let us now give the proof of Theorem 2.4. We write $\chi = k_{U+V}$. The proof imitates the patterns of the proof of (2) $\Rightarrow$ (3) and aims to show that there exists some λ in Λ defined by (2.1) such that, on some open subset of E,

$$\chi''(\theta) = \mathbb{P}(\chi'(\theta))/\lambda, \tag{4.3}$$

thus showing that the variance function of the NEF generated by the law of $U + V$ is $V(m) = \mathbb{P}(m)/\lambda$ on its domain of the means; that is, $U + V$ is Wishart distributed with shape parameter λ (since the domain of existence $-E_+$ of the Laplace transform of the law of $U + V$ is open, any probability generating the NEF must belong to the NEF).

Rao's theorem (Proposition 2.8) and (2.9) imply that, for $-\theta$ in E_+, one has

$$k_U = a\chi \tag{4.4}$$

[note that U concentrated on $\overline{E_+}$ implies that $\Theta(U) \supset -E_+$].

Now, from (2.8) and (4.4), we obtain

$$\begin{aligned}(a\chi'' &+ a^2\chi' \otimes \chi')L_{U+V} \\ &= \mathbb{E}(U \otimes U \exp\langle \theta, U + V\rangle) \\ &= \left[b_{11}(\chi'' + \chi' \otimes \chi') + b_{12}\left(\mathbb{P}\left(\frac{d}{d\theta}\right)\chi + \mathbb{P}(\chi')\right)\right]L_{U+V}.\end{aligned} \tag{4.5}$$

Similarly,

$$\begin{aligned}\bigg(a\mathbb{P}\left(\frac{d}{d\theta}\right)\chi &+ a^2\mathbb{P}(\chi')\bigg)L_{U+V} \\ &= \mathbb{E}(\mathbb{P}(U)\exp\langle \theta, U + V\rangle) \\ &= \left[b_{21}(\chi'' + \chi' \otimes \chi') + b_{22}\left(\mathbb{P}\left(\frac{d}{d\theta}\right)\chi + \mathbb{P}(\chi')\right)\right]L_{U+V}.\end{aligned} \tag{4.6}$$

If we eliminate $\mathbb{P}(d/d\theta)\chi$ between (4.5) and (4.6), we get

$$\begin{aligned}(a^2 &- a\,\mathrm{trace}\,b + \det b)\chi'' \\ &= b_{12}a(1 - a)\mathbb{P}(\chi') + \left(-a^3 + a^2 b_{22} + ab_{11} - \det b\right)\chi' \otimes \chi'.\end{aligned} \tag{4.7}$$

We note that the coefficient of χ'' in (4.7) is the characteristic polynomial of the matrix b evaluated at a. Let us show that b_{12} is not 0 and that a is not an eigenvalue of b. If $b_{12} = 0$, then (4.5) implies

$$(4.8) \qquad (a - b_{11})\chi'' = \left(b_{11} - a^2\right)\chi' \otimes \chi'.$$

But since $0 < a < 1$ (see Proposition 2.8), $(a - b_{11}) = 0$ would imply $\chi' \otimes \chi' = 0$ and $U + V$ would be Dirac: this is impossible. Hence (4.8) would imply $\chi'' = \mu\chi' \otimes \chi'$, with $\mu = (b_{11} - a^2)/(a - b_{11})$. This, in turn, would imply that χ'' has rank 1 everywhere and therefore that $U + V$ is concentrated on a line; this contradicts the hypothesis of Theorem 2.4. Therefore b_{12} is not 0.

Assume now that a is an eigenvalue of b. From (4.7), since $b_{12} \neq 0$, we get that there exists a constant μ such that

$$(4.9) \qquad \mathbb{P}(\chi'(\theta)) = \mu\chi'(\theta) \otimes \chi'(\theta).$$

From the second part of Proposition 2.1 and from (4.9), we see that $\chi'(\theta)$ is a multiple of some primitive idempotent $c(\theta)$. Now, primitive idempotents are on the extremal lines of the convex cone $\overline{E_+}$. Since $\chi'(\theta)$ is the expectation of some element of the NEF generated by the distribution of $U + V$, this implies that $U + V$ is concentrated on an extremal line of $\overline{E_+}$, and this contradicts the hypothesis of Theorem 2.4. Therefore a cannot be an eigenvalue of b.

We have just proved that the coefficients of χ'' and $\mathbb{P}(\chi')$ in (4.7) cannot be 0. Therefore from (4.7) it follows that there exist λ in $\mathbb{R} \setminus \{0\}$ and β in $\mathbb{R}$ such that, for $-\theta$ in E_+,

$$(4.10) \qquad \chi''(\theta) = \mathbb{P}(\chi'(\theta))/\lambda + \beta\chi'(\theta) \otimes \chi'(\theta).$$

It has been proved in Casalis and Letac (1996) [see discussion following (6.17)] that (4.10) implies that the distribution of $U + V$ generates a genuine NEF (i.e., not concentrated on some affine hyperplane) and that $\beta = 0$. Thus (4.3) and Theorem 2.4 are proved.

5. A proof of Olkin and Rubin's theorem. In this section, we will prove the following Proposition 5.1. From this proposition and Theorem 2.4, we obtain a proof of Olkin and Rubin's theorem, as expressed in the form given in Casalis and Letac (1996), Theorem 3.2. The proof of our Theorem 2.4 borrows a number of features from Casalis and Letac (1996) [essentially Proposition 6.1, for our Proposition 2.1; Lemma 6.3, for our Proposition 3.1; and the tedious discussion following (6.17)]. Furthermore, the proof of Proposition 5.1 will use Lemma 5.1 and Proposition 6.2 of Casalis and Letac (1996). However, the present approach is basically simpler and seems to be the normal route, as mountain climbers say, toward Olkin and Rubin's result. This elegant peak has not been very much visited after its discovery: The Science Citation Index acknowledges only three quotations before 1994, and one of them is confusing it with another paper by the same authors and in the same journal, but published in 1964. Srivastava (1965) says that the extension of the result to complex Wishart laws is "under investigation." This aim seems to have been reached only later, in Casalis and Letac (1996).

PROPOSITION 5.1. *Let U and V be two independent random variables concentrated on $\overline{E_+}$ such that $U+V$ belongs to E_+ almost surely and is not concentrated on some half line. Let $g: E_+ \to G$ be a division algorithm, that is, a measurable map such that $g(x)(x) = e$ for all x in E_+. Finally, assume that*

$$Z = g(U+V)(U)$$

is independent of $U+V$ and that the distribution of Z is invariant by K.

Then there exist a in $\mathbb{R}$ and a $(2,2)$ real matrix b such that $\mathbb{E}(U \mid U+V) = a(U+V)$ and (2.8) holds.

PROOF. Since the distribution of Z is K-invariant, $\mathbb{E}(Z)$ is K-invariant. Then, from Lemma 5.1 of Casalis and Letac (1996), we know that there exists a in $\mathbb{R}$ such that $ae = \mathbb{E}(Z)$. Hence

$$ae = \mathbb{E}(g(U+V)(U) \mid U+V) = g(U+V)(\mathbb{E}(U \mid U+V)).$$

Since $g(x)(x) = e$ for all x, we have $(g(x))^{-1}(e) = x$. Hence $a(U+V) = \mathbb{E}(U \mid U+V)$.

Similarly, for (α, β) in $\mathbb{R}^2$, $Q(x) = \alpha x \otimes x + \beta \mathbb{P}(x)$ defines a G-invariant quadratic polynomial taking its values in $L_s(E)$. Hence $\mathbb{E}(Q(Z))$ is a K-invariant element of $L_s(E)$ and, from Proposition 6.1 of Casalis and Letac (1996), there exist λ and μ in $\mathbb{R}$ such that

$$\lambda id_E + \mu e \otimes e = \mathbb{E}(Q(Z)).$$

Thus

$$\begin{aligned}\lambda id_E + \mu e \otimes e &= \mathbb{E}(g(U+V)Q(U)(g(U+V))^* \mid U+V)\\ &= g(U+V)\mathbb{E}(Q(U) \mid U+V)(g(U+V))^*.\end{aligned}$$

Using $(g(U+V))^{-1}(e) = U+V$ again, we get

$$\mathbb{E}(Q(U) \mid U+V) = \lambda \mathbb{P}(U+V) + \mu(U+V) \otimes (U+V), \tag{5.1}$$

since $(g(U+V))^{-1}((g(U+V))^{-1})^* = \mathbb{P}(U+V)$, by applying $\mathbb{P}(gx) = g\mathbb{P}(x)g^*$ [see (3.8) in Casalis and Letac (1996)] to $x = e$ and $g = (g(U+V))^{-1}$. Since λ and μ are linear functions of (α, β), (5.1) proves that there exists b such that (2.8) holds, and the proof is complete. □

6. Proof of Theorem 2.5. The proof of Theorem 2.5 relies on the following proposition:

PROPOSITION 6.1. *Let U be a random variable on E with Wishart distribution $\gamma_{p,\sigma}$. Then $\mathbb{E}(U^{-1})$ is finite if and only if $p > n/r = 1 + d(r-1)/2$. Furthermore, for $p > n/r$ one has*

$$\mathbb{E}(U^{-1}) = \frac{\sigma^{-1}}{p - n/r}, \tag{6.1}$$

$$\mathbb{E}(U^{-1} \otimes U) = \frac{1}{p - n/r}\left(-id_E + p\sigma^{-1} \otimes \sigma\right). \tag{6.2}$$

REMARKS. Equation (6.1) is well known from symmetric real matrices [see Muirhead (1982), pages 97 and 113; note that on page 113, 3.6(b), V^{-1} should be replaced by V]. However, the proof is not that easy, and ours is rather different from the one given by Muirhead. Note also that $p > 1$ is the condition for a gamma random variable U in order that $\mathbb{E}(1/U)$ is finite. Finally, (6.1) and (6.2) imply that, as in Corollary 2.6,

$$\mathrm{Cov}(U^{-1}, U) = \frac{-1}{p - n/r}\, id_E.$$

PROOF. Suppose that $p > n/r$. For x in $\overline{E_+}$, denote

$$g(x) = (\Gamma_E(p))^{-1} (\det x)^{p - n/r},$$

where

$$\Gamma_E(p) = (2\pi)^{(n-r)/2} \Gamma(p) \Gamma(p - d/2) \cdots \Gamma(p - d(r-1)/2). \tag{6.3}$$

Then the density of U with respect to the Lebesgue measure on the Euclidean space E (and restricted to E_+) is

$$\exp - \langle \sigma^{-1}, x \rangle g(x) (\det \sigma)^{-p}. \tag{6.4}$$

A way to see that $\mathbb{E}(U^{-1})$ exists is to observe that, for x in E_+, $x^{-1} \det x$ is actually a polynomial taking its values in E. This comes from the very definition of the determinant of x by the equality

$$x^r - x^{r-1} \operatorname{trace} x + \cdots + (-1)^r e \det x = 0.$$

Since $p > n/r$, this polynomial divided by $\det x$ is integrable on E_+ with respect to $\exp - \langle \sigma^{-1}, x \rangle g(x)\, dx$.

Let us assume that θ belongs to $-E_+$. Since E is Euclidean, we shall write the differential of a real differentiable function on E as a gradient, that is, as an element of E. Clearly, the function

$$x \mapsto \exp\langle \theta, x \rangle g(x)$$

has differential

$$x \mapsto \exp\langle \theta, x \rangle g(x) \big((p - n/r) x^{-1} + \theta \big),$$

since the differential of $x \mapsto \log \det x$ is $x \mapsto x^{-1}$. Furthermore, this function is 0 on the boundary $\overline{E_+} \setminus E_+$. An application of Stokes formula then gives

$$\int_{E_+} \exp\langle \theta, x \rangle g(x) \big((p - n/r) x^{-1} + \theta \big)\, dx = 0. \tag{6.5}$$

For $\theta = -\sigma^{-1}$ and given the density (6.4), (6.5) gives now (6.1). Taking the differential of the first member of (6.5) with respect to θ yields

$$\int_{E_+} \exp\langle \theta, x \rangle g(x) \big((p - n/r) x^{-1} \otimes x + \theta \otimes x + id_E \big)\, dx = 0.$$

Again, for $\theta = -\sigma^{-1}$ and given the density (6.4), the preceding equality and

$$\int_{E_+} \exp - \langle \sigma^{-1}, x \rangle g(x)(\det \sigma)^{-p} x\, dx = p\sigma,$$

which is the expectation of U, yield (6.2).

Finally, from (6.1), we get

$$\begin{aligned} &\int_{E_+} \exp\langle -\sigma^{-1}, x\rangle (\det x)^{p-n/r} \operatorname{trace}(x^{-1})\, dx \\ &\qquad = \frac{\operatorname{trace}(\sigma^{-1})}{p - n/r} (\det \sigma)^p \Gamma_E(p), \end{aligned} \tag{6.6}$$

which, as a function of p, is a Laplace transform of a positive function. The second member of (6.6) has a pole at $p = n/r$, and, from (6.3), we see that the integral does not exist for $p \leq n/r$. The proof of Proposition 6.1 is complete. □

Proof of Theorem 2.5. We apply the principle of equality (3.9) to the pair $Q(u) = u^{-1}$ and $Q_1 = cQ$, with c defined by (2.10). Replacing σ in (6.1) by $\sigma_\theta = (\sigma^{-1} - \theta)^{-1}$ will give the first of the two equalities (2.11). One proceeds in the same way with (6.2), to get the second equality in (2.11), by dealing with the pair $Q(u) = u^{-1} \otimes u$ and $Q_1 = c_1 id_E + c_2 Q$, where c_1 and c_2 are defined by (2.10). The proof of Theorem 2.5 is then complete. □

To end this section, we observe that an analogue of Proposition 2.1 for the inverse is available here. We skip its proof, which imitates the first part of Proposition 2.1. Here again, the symmetry of $Q(e)$ is important.

Proposition 6.2. *Let $\mathscr{Q}_1$ be the linear space of functions Q: $E_+ \to L(E)$ such that for all g in G one has*

$$Q(g(x)) = g^{*-1} Q(x) g^* \tag{6.7}$$

and such that $Q(e)$ is in $L_s(E)$. Then $\mathscr{Q}_1$ has dimension 2, *and the pair $x \mapsto x^{-1} \otimes x$ and $x \mapsto id_E$ defines a basis of $\mathscr{Q}_1$.*

7. Proof of Theorem 2.7. Since (2.9) is true, (2.12) and (4.1) hold. Recall that $0 < a < 1$. For θ in $-E_+$, we denote $h_U(\theta) = \mathbb{E}(U^{-1} \exp\langle \alpha, U\rangle)$. Then the first equality of (2.11) implies that

$$h_U(\theta) L_V(\theta) = c h_{U+V}(\theta). \tag{7.1}$$

Taking the differential of both sides of (7.1), we get

$$h'_U(\theta) L_V(\theta) + h_U(\theta) \otimes L'_V(\theta) = c h'_{U+V}(\theta). \tag{7.2}$$

The second equality of (2.11) implies that

$$h'_U(\theta) L_V(\theta) = c_1 L_{U+V}(\theta) id_E + c_2 h'_{U+V}(\theta). \tag{7.3}$$

Eliminating $h'_{U+V}(\theta)$ between (7.2) and (7.3) and using (4.1), we get

(7.4) $(c_2 - c)h'_U(\theta) + c_2(1-a)h_U(\theta) \otimes \chi'(\theta) + cc_1 L_U(\theta) id_E = 0.$

Observe that (2.11) implies that $c > 0$, since U^{-1} and $(U+V)^{-1}$ are in E_+. Having $c_2 - c = 0$ in (7.4) implies

$$h_U(\theta) \otimes \chi'(\theta) = -\frac{c_1}{1-a} L_U(\theta) id_E,$$

and this is impossible since the left-hand side of the equation has rank 1 and the right-hand side has rank 0 if $c_1 = 0$ and rank n if $c_1 \neq 0$. Thus $c_2 - c \neq 0$.

We adapt to (7.4) the method of resolution of one-dimensional linear and nonhomogeneous ordinary differential equations, by considering the function w on $-E_+$, taking its values in E_+ and defined by $h_U = L_U^{\alpha} w$, where

$$\alpha = -\frac{c_2(1-a)}{(c_2 - c)a}.$$

Then (7.4) becomes

$$w' = \frac{cc_1}{c - c_2} L_U^{1-\alpha} id_E. \tag{7.5}$$

We now prove that the second member of (7.5), denoted by ψid_E, must be a constant with respect to θ on $-E_+$. This comes from the fact that w'' is a Hessian and must be symmetric; that is, the function on E^2 defined by

$$(h,k) \mapsto w''(\theta)(h,k) = \langle \psi'(\theta), h\rangle k$$

is symmetric in h and k. Clearly, this implies that $\psi'(\theta) = 0$ on $-E_+$. Since $-E_+$ is connected, ψ is a real constant and is equal to 0 if and only if $c_1 = 0$. In this case, from (7.5), it follows that w is a constant belonging to E. Thus, for any h in E orthogonal to w, we have $\mathbb{E}(\langle U^{-1}, h\rangle \exp\langle \theta, U\rangle) = 0$ for all θ: This implies that U^{-1} is concentrated on $\mathbb{R}w$, which implies, in turn, that U is concentrated on $\mathbb{R}w^{-1}$. Proposition 2.8 implies that the same is true for $U + V$, and this contradicts the hypothesis of Theorem 2.7. Thus $\psi \neq 0$, and α must be equal to 1, which is equivalent to $c_2 = ac$. Thus (7.5) becomes $w'(\theta) = c_1/(1-a) id_E$, and there exists a constant c_0 in E such that $w(\theta) = c_1/(1-a)(\theta - c_0)$. Since $h_U = L_U w$ is in E_+ for θ in $-E_+$, this implies that $c_1 < 0$ and that c_0 is in E_+. We write $c_0 = \sigma^{-1}$ and introduce the positive numbers p and q defined by $p - n/r = -(1-a)/c_1$ and $a(p+q) = p$. So, for θ in $-E_+$ we now have

$$h_U(\theta) = \frac{\sigma^{-1} - \theta}{p - n/r} L_U(\theta), \tag{7.6}$$

and we are going to deduce from this that the distribution of U is $\gamma_{p,\sigma}$. We introduce the positive measure μ on E_+ such that the distribution of U is

$$\exp -\langle \sigma^{-1}, x\rangle g(x)\mu(dx), \tag{7.7}$$

where g has been defined in the proof of Proposition 6.1. We want to show that μ is proportional to the Lebesgue measure.

Using density (7.7) for computing h_U in (7.6) and replacing $\theta - \sigma^{-1}$ by θ in $-E_+$, we obtain that

$$\int_{E_+} (\exp\langle \theta, x\rangle g(x))' \mu(dx) = 0. \tag{7.8}$$

Since $\exp\langle \theta, x\rangle g(x)$, as a function of x, is 0 on the boundary of E_+, using the Schwartz derivative $(\mu)'$, we see that (7.8) implies

$$-\int_{E_+} \exp\langle \theta, x\rangle g(x)(\mu)'(dx) = 0$$

for all θ in $-E_+$. Thus $g(x)(\mu)'(dx) = 0$, and μ is proportional to the Lebesgue measure. The proof is now complete. □

8. Further comments. Recall the following result by Huang, Li and Lo Huang (1994):

THEOREM. *Let U and V be positive, nondegenerate random variables such that $\mathbb{E}(U^{-2})$ is finite. Then there exists a and $b > 0$ such that*

$$\mathbb{E}(U^{-1} \mid U + V) = a(U+V)^{-1}, \qquad \mathbb{E}(U^{-2} \mid U + V) = b(U+V)^{-2}$$

if and only if there exist $p > 2$, $q > 0$ and $\sigma > 0$ such that U and V have distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$.

[Huang, Li and Lo Huang add the hypothesis $\mathbb{E}(V^2) < +\infty$. This is not necessary if one proceeds as follows for the proof: Define for $\theta \le 0$ the function $f_U(\theta) = \mathbb{E}(U^{-2}\exp(\theta U))$. One gets easily

$$f'_U L_V = a f'_{U+V}, \tag{8.1}$$

$$f_U L_V = b f_{U+V}. \tag{8.2}$$

Derive (8.2) and eliminate f'_{U+V} with (8.1); then derive (8.1) and use $f''_{U+V} = f''_U L_V$. One gets the two equations

$$\left(1 - \frac{b}{a}\right) f'_U L_V + f_U L'_V = 0, \qquad (1-a) f''_U L_V + f'_U L'_V = 0,$$

and the elimination of (L_V, L'_V) between them leads to a differential equation for f_U.]

Now, a referee has asked for the generalization to the Wishart distributions of the above result. It seems to be an interesting and difficult problem which needs new methods. As a partial answer, we shall content ourselves here to mention the following direct result:

THEOREM 8.1. *Let U and V be independent random variables with Wishart distributions $\gamma_{p,\sigma}$ and $\gamma_{q,\sigma}$ on the simple Euclidean Jordan algebra E, with Pierce constant d and rank r. Assume that $A_1 = p - 1 - d(r-1)/2 > 0$. Denote also $A_2 = p - 2 - d(r-1)/2$, $B = p - 1 - d(r-2)/2$ and consider* *the matrices*

$$B(p) = \frac{1}{A_1 A_2 B} \begin{bmatrix} A_2 + \dfrac{d}{2} & 1 \\ \dfrac{d}{2} & A_1 \end{bmatrix}$$

and

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = B(p) B(p+q)^{-1}.$$

Then

$$\mathbb{E}(U^{-1} \otimes U^{-1} \mid U + V) = b_{11}(U+V)^{-1} \otimes (U+V)^{-1} + b_{12}\mathbb{P}\big((U+V)^{-1}\big), \tag{8.3}$$

$$\mathbb{E}\big(\mathbb{P}(U^{-1})\big) = b_{21}(U+V)^{-1} \otimes (U+V)^{-1} + b_{22}\mathbb{P}\big((U+V)^{-1}\big). \tag{8.4}$$

A statement similar to Corollary 2.3 also holds by considering the line eigenvectors of the matrix $B(p)$. The proof of Theorem 8.1 is more difficult than the proofs of Theorems 2.2 and 2.5. It relies on the extension to the Jordan algebras of the formulas of von Rosen (1988), Theorem 3.1(ii), and Das Gupta (1968), Lemma 2.4(ii). The tool for this is a result on the Peirce decomposition of a Wishart variable obtained by Massam and Neher (1997), Theorem 4.3.1. The proof of our Theorem 8.1 can be found in Letac and Massam (1997), which contains a general study of equalities of the type

$$\mathbb{E}\Big(c_p^{-1}(Q)(U) \mid U + V\Big) = c_{p+q}^{-1}(Q)(U+V),$$

where U and V are independent and have distribution in the same exponential dispersion model (e.g., the Wishart one), where Q belongs to some finite-dimensional linear space ν of functions and where $(c_p;\ p \in \Lambda)$ is a suitable family of automorphisms of ν.

Acknowledgment. We thank Muriel Casalis for several comments.

REFERENCES

CASALIS, M. (1990). Familles exponentielles naturelles invariantes par un groupe. Thèse, Univ. Paul Sabatier.

CASALIS, M. (1992). Un estimateur de la variance pour une famille exponentielle à fonction variance quadratique. *C.R. Acad. Sci. Paris Sér. I, Math.* **314** 143–146.

CASALIS, M. and LETAC, G. (1996). The Lukacs–Olkin–Rubin characterization of the Wishart distributions on symmetric cones. *Ann. Statist.* **24** 763–786.

DAS GUPTA, S. (1968). Some aspects of discrimination function coefficients. *Sankhyā Ser. A* **30** 387–400.

FARAUT, J. and KORANYI, A. (1994). *Analysis on Symmetric Cones*. Oxford Univ. Press.

HUANG, W.-J., LI, S.-H. and HUANG, M.-N. L. (1994). Characterizations of the Poisson process as a renewal process via two conditional moments. *Ann. Inst. Statist. Math.* **46** 351–360.

KAWATA, Y. (1972). *Fourier Analysis and Probability*. Wiley, New York.

LETAC, G. and MASSAM, H. (1997). Representations of an exponential dispersion model (with an appendix by E. Neher). Technical report, Univ. Paul Sabatier.

LUKACS, E. (1955). A characterization of the gamma distribution. *Ann. Math. Statist.* **26** 319–324.

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. (1997). On transformations and determinants of Wishart variables on symmetric cones. *J. Theoret. Probab.* **10** 867–902.

MUIRHEAD, R. (1982). *Aspects of Multivariate Analysis*. Wiley, New York.

OLKIN, I. and RUBIN, H. (1962). A characterization of the Wishart distribution. *Ann. Math. Statist.* **33** 1272–1280.

RAO, C. R. (1948). On a problem of Ragnar Frish. *Econometrica* **15** 245–249.

SRIVASTAVA, M. S. (1965). On complex Wishart distribution. *Ann. Math. Statist.* **36** 313–317.

VON ROSEN, D. (1988). Moments for the inverted Wishart distribution. *Scand. J. Statist.* **15** 97–109.

WANG, F. (1981). Extensions of Lukacs' characterization of the gamma distribution. *Analytic Methods in Probability Theory. Lecture Notes in Math.* **861** 166–177. Springer, New York.

WESOŁOWSKI, J. (1990). A constant regression characterization of the gamma law. *Adv. in Appl. Probab.* **22** 488–489.

LABORATOIRE DE STATISTIQUE
ET PROBABILITÉS
UNIVERSITÉ PAUL SABATIER
31062 TOULOUSE CEDEX
FRANCE
E-MAIL: letac@cict.fr

DEPARTMENT OF MATHEMATICS
AND STATISTICS
YORK UNIVERSITY
NORTH YORK, ONTARIO M3J 1P3
CANADA
AND
DEPARTMENT OF STATISTICS
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22903