Bulletin
de la SOCIÉTÉ MATHÉMATIQUE DE FRANCE

WHY JORDAN ALGEBRAS ARE NATURAL IN STATISTICS: QUADRATIC REGRESSION IMPLIES WISHART DISTRIBUTIONS

G. Letac & J. Wesołowski

Tome 139
Fascicule 1

2011

SOCIÉTÉ MATHÉMATIQUE DE FRANCE
Publié avec le concours du Centre national de la recherche scientifique
pages 129-144

Bull. Soc. math. France
139 (1), 2011, p. 129–144

WHY JORDAN ALGEBRAS ARE NATURAL IN STATISTICS: QUADRATIC REGRESSION IMPLIES WISHART DISTRIBUTIONS

BY G. LETAC & J. WESOŁOWSKI

ABSTRACT. — If the space Q of quadratic forms in $\mathbb{R}^n$ is splitted in a direct sum $Q_1 \oplus \cdots \oplus Q_k$ and if X and Y are independent random variables of $\mathbb{R}^n$, assume that there exist a real number a such that $E(X|X+Y) = a(X+Y)$ and real distinct numbers $b_1, ..., b_k$ such that $E(q(X)|X+Y) = b_i q(X+Y)$ for any q in Q_i. We prove that this happens only when $k = 2$, when $\mathbb{R}^n$ can be structured in a Euclidean Jordan algebra and when X and Y have Wishart distributions corresponding to this structure.

RÉSUMÉ (*Pourquoi les algèbres de Jordan sont-elles naturelles en statistiques? La régression quadratique implique la distribution de Wishart*)

Si l'espace Q des formes quadratiques sur $\mathbb{R}^n$ est décomposé en une somme directe $Q_1 \oplus \cdots \oplus Q_k$ et si X et Y sont des variables aléatoires indépendantes de $\mathbb{R}^n$, supposons qu'il existe un nombre réel a tel que $E(X|X+Y) = a(X+Y)$ ainsi que des nombres réels distincts $b_1, ..., b_k$ tels que $E(q(X)|X+Y) = b_i q(X+Y)$ pour tout q de Q_i. Nous montrons que cela n'arrive que pour $k = 2$, que lorsque $\mathbb{R}^n$ peut être structuré en algèbre de Jordan euclidienne et que lorsque X et Y suivent des lois de Wishart correspondant à cette structure.

Texte reçu le 2 juin 2009, accepté le 6 novembre 2009

G. LETAC, Laboratoire de Statistique et Probabilités, Université Paul Sabatier, 31062 Toulouse, France • *E-mail :* `letac@cict.fr`

J. WESOŁOWSKI, Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Warszawa, Poland • *E-mail :* `wesolo@mini.pw.edu.pl`

2000 Mathematics Subject Classification. — 60H10, 62H05.

Key words and phrases. — Symmetric cones, random matrices, characterization of Wishart laws.

BULLETIN DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE 0037-9484/2011/129/$ 5.00

I. Introduction

Let S_r be the set of (r,r) real symmetric matrices and let X and Y be independent random variables valued in S_r such that they are Wishart distributed $\gamma_{p,\sigma}$ and $\gamma_{p',\sigma}$, which means that

$$\mathbb{E}(e^{-\mathrm{tr}\,\theta X}) = \det(I_r + \theta\sigma)^{-p} \tag{1.1}$$

where θ and σ are in the set P_r of the positive definite elements of S_r and p is in

$$\Lambda = \{\frac{1}{2}, \dots, \frac{r-1}{2}\} \cup (\frac{r-1}{2}, \infty) \tag{1.2}$$

(In (1.1) tr means trace). Note that for $a = p/(p+p')$

$$\mathbb{E}(X|X+Y) = a(X+Y). \tag{1.3}$$

Assume furthermore that $p+p' > \frac{r-1}{2}$. This implies that $(X+Y)^{-1}$ exists. Then it is known that $Z = (X+Y)^{-1/2}X(X+Y)^{-1/2}$ and $X+Y$ are independent and that $Z \sim uZu^T$ for any orthogonal (r,r) matrix u. There are many consequences, nuances and characterizations of the Wishart distributions related to this result. One of these consequences is the following fact: for any $s \in S_r$ consider the two quadratic forms on S_r defined by

$$q_1^s(x) = \frac{1}{2}\mathrm{tr}^2(xs) + \mathrm{tr}\,(sxsx), \;\; q_2^s(x) = \mathrm{tr}^2(xs) - \mathrm{tr}\,(sxsx) \tag{1.4}$$

and the two numbers

$$b_1 = \frac{p}{p+p'}\,\frac{p+1}{p+p'+1}, \;\; b_2 = \frac{p}{p+p'}\,\frac{p-\frac{1}{2}}{p+p'-\frac{1}{2}}.$$

Then for $i = 1,2$ and for any s

$$\mathbb{E}(q_i^s(X)|X+Y) = b_i q_i^s(X+Y) \tag{1.5}$$

This is the particular case $d = 1$ of Corollary 2.3 of Letac and Massam (1998). An important fact about this set $(q_1^s, q_2^s)_{s\in S_r}$ is that it spans the whole space of quadratic forms $\mathcal{Q}$ on S_r (since if $q^s(x) = \mathrm{tr}^2(xs)$ then $\{q^s\ ;\ s \in S_r\}$ spans $\mathcal{Q}$). More specifically denote by $\mathcal{Q}_i$ the subspace of $\mathcal{Q}$ generated by $\{q_i^s\ ;\ s \in S_r\}$. Then $\mathcal{Q} = \mathcal{Q}_1 \oplus \mathcal{Q}_2$ (see for instance Theorem 5.2 below for a proof).

The aim of the paper is to prove a reciprocal statement of (1.3) and (1.5): Let V be a linear real finite dimensional space (instead of S_r) and denote by $\mathcal{Q}$ the space of all quadratic forms on V. Fix a decomposition $\mathcal{Q} = \mathcal{Q}_1 \oplus \mathcal{Q}_2 \oplus \cdots \oplus \mathcal{Q}_k$ with $k \geq 2$ as a direct sum of linear subspaces. Consider two independent random variables X and Y with exponential moments satisfying (1.3) for some a and $\mathbb{E}(q(X)|X+Y) = b_i q(X+Y)$ for all $q \in \mathcal{Q}_i$ and for some distinct real numbers $b_1, \dots, b_k$. We show that under these circumstances, necessarily $k = 2$ and X and Y are Wishart distributed in the following sense: there necessarily exists a structure of Euclidean Jordan algebra on V (like symmetric matrices, Hermitian matrices, or space with a Lorentz cone) such that X and Y are Wishart on the symmetric cone associated to it. Section 5 contains more detailed information about the two spaces Q_1 and Q_2 of quadratic forms on S_r (or more generally, on a Euclidean Jordan algebra)

II. Some history of the subject

WISHART DISTRIBUTIONS ON S_r. Wishart distributions have been introduced by J. Wishart (1928) as distributions of $Z_1Z_1^T + \cdots + Z_NZ_N^T \sim \gamma_{N/2,2\Sigma}$ where $Z_1, \ldots, Z_N$ are iid in $\mathbb{R}^r$ such that $Z_i \sim N(0, \Sigma)$. Elegant calculations about them are in Bartlett (1933) and the classical reference is Muirhead (1982). For the space S_r of (r, r) real symmetric matrices the extension of the definition of $\gamma_{p,\sigma}$ from a half integer p to the whole set Λ defined by (1.2) is made in the fundamental paper of Olkin and Rubin (1962). Proving that a distribution $\gamma_{p,\sigma}$ on the semi positive definite matrices such that (1.1) holds only if p is in Λ was considered as a challenge by statisticians (see Eaton (1983)) although the appendix of Olkin and Rubin contains an unnoticed proof of it (and unfortunately erroneous: see Casalis and Letac (1994)). This conjecture was independently proved by Shanbhag (1988) and Peddada and Richards (1989) by quite different means, although a solution already appeared in Gyndikin (1975) and seems to have been well known by analysts, who also call the set Λ and its extensions the Wallach set (see Lassalle (1987) for proofs and references).

Lukacs-Olkin-Rubin Theorem. — Wishart distributions on S_r are the most natural generalization of the gamma distributions on the positive line. Lukacs (1956) shows that if X and Y are positive, independent non Dirac random variables and if $Z = X/(X + Y)$, then Z and $X + Y$ are independent if and only if there exists $\sigma, p, p' > 0$ such that $X \sim \gamma_{p,\sigma}$ and $Y \sim \gamma_{p',\sigma}$. This was extended to S_r by Olkin and Rubin (1962) by a proper definition of Z such that Z is symmetric (for instance by choosing $Z = (X + Y)^{-1/2}X(X + Y)^{-1/2}$ or by choosing $Z = C^{-1}X(C^{-1})^T$ where C is the triangular matrix with positive diagonal elements coming from the Cholesky decomposition $CC^T = X + Y$). They show that if X and Y are independent non Dirac random semi positive definite matrices in S_r such that $X+Y$ is invertible and such that $Z \sim uZu^T$ for any orthogonal (r, r) matrix u then Z and $X + Y$ are independent if and only if there exists a positive definite matrix σ and p and p' in Λ with $p+p' > (r-1)/2$ such that $X \sim \gamma_{p,\sigma}$ and $Y \sim \gamma_{p',\sigma}$. If Z is defined as $(X+Y)^{-1/2}X(X+Y)^{-1/2}$, Bobecka and Wesołowski (2002) have shown that the invariance hypothesis for Z by the orthogonal group can be dropped provided one assumes that X and Y have smooth densities. Removing this assumption of density is still a challenge.

Wishart distributions on Hermitian matrices and on Euclidean Jordan algebras. — Since normal distributions on Hermitian spaces have been considered (see e.g., Goodman (1963)), therefore Wishart distributions on Hermitian matrices occur naturally. Actually physicists considered them quite early (see Mehta (2004)). Carter (1975) in an unpublished PhD thesis extends Olkin and Rubin to this case.

On the other hand, works on the classification of natural exponential families by their variance function have led to the observation that the exponential family $\{\gamma_{p,\sigma}; \sigma \in P_r\}$ of Wishart distributions on S_r with fixed shape parameter $p \in \Lambda$ has a variance function which is the map from S_r into itself $x \mapsto V(m)(x) = \frac{1}{p}mxm$ where m is in P_r. In other terms, this means that if κ is a cumulant function of $\gamma_{p,\sigma}$ then for all x in S_r we have

$$\kappa''(\theta)(x) = \frac{1}{p}\kappa'(\theta)x\kappa'(\theta).$$

Facts about multivariate distributions such that their corresponding variance functions are quadratic in the mean are collected in Letac (1989). In particular, Wishart distributions obtained from simple Euclidean Jordan algebras are described there. An indispensable reference for simple Euclidean Jordan algebras is Faraut and Koranyi (1994) always abreviated F.-K. below. Recall that simple Euclidean Jordan algebras are basically in one to one correspondence with the irreducible symmetric cones (self dual cones in Euclidean space such that the group of automorphisms of the cone acts transitively on it), in the way that S_r is linked to P_r. A quick definition of the Wishart distribution $\gamma_{p,\sigma}$ on the Jordan algebra V with rank r, Peirce constant d, cone $\overline{\Omega}$ of square elements, trace and determinant function tr and det can be done by its Laplace transform

$$\int_{\overline{\Omega}} e^{-\operatorname{tr}\theta x}\gamma_{p,\sigma}(dx) = \det(e+\theta\sigma)^{-p}$$

where σ is in the interior Ω of $\overline{\Omega}$ and where p is in the Gyndikin set of the Jordan algebra V defined by

$$\Lambda_V = \{\frac{d}{2}, d, \dots, \frac{d}{2}(r-1)\} \cup (\frac{d}{2}(r-1), \infty). \tag{2.6}$$

While the definition of determinant is the standard one for S_r and for Hermitian matrices, it requires some care for the three other types of Jordan algebras: quaternionic Hermitian matrices, 27 dimensional Albert algebra and the algebra of the Lorentz cone.

Particular cases of use of Wishart distributions on Jordan algebras in statistics occurred earlier (Andersson (1975) for the Hermitian and quaternionic cases, and Jensen (1988) for the Lorentz cone, with its deep connexions to Clifford algebras). Jordan algebras are the natural framework for Wishart