JOURNAL OF MULTIVARIATE ANALYSIS **63**, 105–118 (1997)
ARTICLE NO. MV971693

Multidimensional Bhattacharyya Matrices and Exponential Families

Denys Pommeret

Université Paul Sabatier, Laboratoire de Statistique et Probabilités, 118, route de Narbonne, 31062 Toulouse cedex, France

Shanbhag (1972, 1979) has characterized the distributions belonging to an exponential family on $\mathbb{R}$ such that the Bhattacharyya matrix is diagonal. Since then, this set of distributions has been classed by Morris (1982) and is referred to as the class of quadratic natural exponential families. In this paper we consider a multidimensional extension of Shanbhag and we obtain a characterization of the quadratic natural exponential families on $\mathbb{R}^d$. © 1997 Academic Press

1. INTRODUCTION

The class of quadratic natural exponential families (*NEF*) on $\mathbb{R}$ is well known and has been described by Morris (1982) as admitting six types of distributions namely the Gaussian, Poisson, gamma, binomial, negative binomial and hyperbolic types. On $\mathbb{R}^d$, the full class of the quadratic *NEF* has not yet been completely determined. However, the subclass of simple quadratic *NEF* has been determined by Casalis (1996) and may be split into $2d+4$ types: the $d+1$ Poisson Gaussian types, the $d+1$ negative multinomial gamma types, the multinomial and the hyperbolic types (see Section 4). Before then, Bhattacharyya (1946) has considered the covariance matrix of $(f^{(1)}(x,m)/f(x,m), ..., f^{(i)}(x,m)/f(x,m), ...)$ where $f(x,m)$ is a m-parametrized density and $f^{(i)}(x,m)$ is the ith derivative of $m \to f(x,m)$. Seth (1949) has proved that this matrix is diagonal when f is one of the six previous distributions on $\mathbb{R}$. Shanbhag (1972, 1979) shows that the distribution assumptions are necessary as well as sufficient for the diagonality. Clearly this last result characterizes the class of the quadratic *NEF*. In addition, Shanbhag has established that the diagonality of the $s \times s$ ($s \geqslant 3$) Bhattacharyya matrix is equivalent to the diagonality of the 3×3 one. The aim of the present paper is to generalize these results to $\mathbb{R}^d$. For that, let

Received June 6, 1996; revised May 15, 1997.

AMS 1991 subject classifications: 62H05.

Key words and phrases: multidimensional Bhattacharyya matrices, multidimensional exponential families, multivariate orthogonal polynomials, variance functions.

0047-259X/97 $25.00

us introduce our framework: the natural exponential families (NEF) and the Bhattacharyya matrices.

Let ν be a probability measure on $\mathbb{R}^d$, with mean m_0, not concentrated on an affine hyperplane and let L_ν be its Laplace transform, i.e., $L_\nu(\theta) = \int_{\mathbb{R}^d} \exp < \theta, x > \nu(dx)$ where $\theta = {}^t(\theta_1, ..., \theta_d)$, $x = {}^t(x_1, ..., x_d)$ and $\langle \theta, x \rangle = \sum_{i=1}^d \theta_i x_i$. We assume that $\Theta(\nu)$, the interior of the set $\{\theta \in \mathbb{R}^d;\ L_\nu(\theta) < +\infty\}$, is not empty. For $\theta \in \Theta(\nu)$ we write $k_\nu(\theta) = log(L_\nu(\theta))$. Then the first derivative, k'_ν, defines a diffeomorphism between $\Theta(\nu)$ and its image M_F (see Letac 1992) and its inverse function is written ψ_ν. The set $F = F(\nu) = \{P(m, F);\ m \in M_F\}$ defines the *natural exponential family* (*NEF*) generated by ν, where the density of $P(m, F)$ with respect to ν is:

$$f_\nu(x, m) = \exp\{\langle \psi_\nu(m), x \rangle - k_\nu(\psi_\nu(m))\}. \tag{1}$$

Since m is the mean of the probability $P(m, F)$ this parametrization is referred to as the *mean parametrization.* For all $m \in M_F$, $V_F(m) = (V_{ij}(m))_{(i,j) \in \mathbb{N}^d}$ denotes the covariance matrix of $P(m, F)$ and the function V_F is called the *variance-function* of F. A *NEF* is called *quadratic* if $V_{ij}(m)$ is a polynomial in m with degree two. It is called *simple quadratic* if the term with degree two in $V_{ij}(m)$ is $am_i m_j$, where a is a real constant independent of (i, j); i.e. $V_F(m) = am\,{}^tm + B(m) + C$ where $B\colon M_F \to \mathbb{R}^{d \times d}$ is linear and C is $d \times d$ symmetric. A particular case is a negative where we necessarily have $a = -1/N$, where N is a positive integer (see Casalis 1996).

Giving an open set $I \subseteq \mathbb{R}^d$ and a C^∞ diffeomorphism $t\colon I \to \Theta(\nu)$ we may consider another parametrization of F and F can be expressed as: $F(\nu) = \{P(t(z), \nu);\ z \in I\}$, where the density of $P(t(z), \nu)$ with respect to ν is

$$h_\nu(x, z) = f_\nu(x, k'_\nu(t(z))) = \exp\{\langle t(z), x \rangle - k_\nu(t(z))\}. \tag{2}$$

We now define the multidimensional Bhattacharyya matrices based on multivariate polynomials. For that we use the following definition: a polynomial P in $x \in \mathbb{R}^d$ of degree $k \in \mathbb{N}$ may then be given by $P(x) = \sum_{q \in \mathbb{N}^d,\ |q| \leqslant k} \alpha_q x^q$, where $x^q = x_1^{q_1} \cdots x_d^{q_d}$, $|q| = q_1 + \cdots + q_d$ and at least one of the real numbers α_q is non zero when $|q| = k$. For all $n \in \mathbb{N}^d$ and for all $A \in GL(R^d)$ (the set of the $d \times d$ invertible matrices), on $\mathbb{R}^d \times I$ let

$$L_{A,n}(x, z) = \frac{h_\nu^{(n)}(x, z)(Ae_1, ..., Ae_d)}{h_\nu(x, z)}, \tag{3}$$

where $e_1, ..., e_d$ denotes the canonical basis of $\mathbb{R}^d$, and $h_\nu^{(n)}(x, z)$ $(Ae_1, ..., Ae_d)$ is the $|n| = n_1 + n_2 + \cdots + n_d$ derivative of $z \mapsto h_\nu(x, z)$ in the $|n|$ directions Ae_1 (n_1 times), ..., Ae_d (n_d times). We obtain the following properties:

LEMMA 1.1. (i) *For any* $(A, z) \in GL(\mathbb{R}^d) \times I$, $L_{A,n}(x, z)$ *is a polynomial in* x *of degree* $|n|$. *Furthermore, the* $(L_{A,n})_{n \in \mathbb{N}^d}$ *form a basis of the set of all the polynomials on* $\mathbb{R}^d$.

(ii) *For all* $z_0 \in I$, *there exists* $r > 0$ *such that if* $z \in B(z_0, r)$ (*i.e.,* $|z_i - z_{0_i}| < r$, $\forall i = 1 \cdots d$), *then*

$$\sum_{n \in \mathbb{N}^d} \frac{(z - A^{-1} z_0)^n}{n!} L_{A,n}(x, z_0) = \frac{h_\nu(x, Az)}{h_\nu(x, z_0)}.$$

(*here, if* $\sigma = (\sigma_1, ..., \sigma_d)$ *is in* $\mathbb{R}^d$ *and* $n = (n_1, ..., n_d)$ *is in* $\mathbb{N}^d$, *we write* $\sigma^n/n!$ *for* $\prod_{i=1}^d \sigma_i^{n_i}/n_i!$).

Proof. Part (i) is shown by induction on $|n|$ and Part (ii) comes from the analyticity of ψ_ν and k_ν. ∎

For all $z \in I$ and $(n, m) \in (\mathbb{N}^d)^2$, we denote

$$J_{A;nm}(z) = \int_{\mathbb{R}^d} L_{A,n}(x, z)\, L_{A,m}(x, z)\, P(t(z), \nu)(dx). \tag{4}$$

The infinite matrix

$$J_A(z) = (J_{A;nm}(z))_{(n,m) \in (\mathbb{N}^d)^2}, \tag{5}$$

defines the *Bhattacharyya matrix*. For all $(k, l) \in (\mathbb{N}^d)^2$ we consider the $\binom{d+k-1}{k} \times \binom{d+l-1}{l}$ submatrices of $J_A(z)$ (where $\binom{a}{b} = a!/b!(a-b)!$):

$$J_A^{k,l}(z) = (J_{A;nm}(z))_{n,m;\,|n|=k,\,|m|=l}, \tag{6}$$

and we say that the Bhattacharyya matrix $J_A(z)$ is *pseudo-diagonal* if for all $k, l \in \mathbb{N}$, $J_A^{k,l}(z) = 0$ when $k \neq l$.

We obtain the following two results: $J_A(z)$ is diagonal (resp. pseudo-diagonal) if and only if F is simple quadratic (resp. quadratic).

Clearly, for $d = 1$, the two results are the same; Shanbhag has established them in (1972) and (1979). Furthermore, in these papers Shanbhag performs after Meixner (1934) the classification of the quadratic *NEF* on $\mathbb{R}$ (as redone later on by C. Morris (1982)). Actually this classification is an essential tool for his proof, a part of which being done by investigation of the six types of quadratic *NEF*.

In $\mathbb{R}^d$, such a classfication of the quadratic *NEF* is only partially available. For this reason, to prove the main results of this article (Theorems 3.1 and 3.2), we have to rely on substitutes, which are our Theorems 2.1 and 2.2 in Section 2, which give generalizations of (respectively) Meixner (1934) and Feinsilver (1986). Finally Section 4 provides the calculus of the multidimensional diagonal Bhattacharyya matrices.

2. ORTHOGONAL POLYNOMIALS

Here we consider a particular case of the polynomials $L_{A,n}$ that is: for $(A, n) \in GL(\mathbb{R}^d) \times \mathbb{N}^d$, we define on $\mathbb{R}^d$

$$P_{A,n}(x) = L_{A,n}(x, m_0) = f_\nu^{(n)}(x, m_0)(Ae_1, ..., Ae_d), \tag{7}$$

where ν is the probability measure on $\mathbb{R}^d$ with mean m_0 given in Section 1. Thus, by Lemma 1.1, $P_{A,n}(x)$ is a polynomial of degree $|n|$ and the $(P_{A,n})_{n \in \mathbb{N}^d}$ form a basis of the space of all polynomials on $\mathbb{R}^d$.

We say that the $(P_{A,n})_{n \in \mathbb{N}^d}$ are *ν-orthogonal* (resp. *ν-pseudo-orthogonal*) if $\int P_{A,n}(x) P_{A,m}(x)\, \nu(dx) = 0$ when $n \neq m$ (resp. $|n| \neq |m|$). We characterize the polynomials $P_{A,n}$ among all the sequences of ν-(pseudo-) orthogonal polynomials by a condition equivalent to Meixner's (1934).

THEOREM 2.1. *Let F be a NEF on $\mathbb{R}^d$ and ν the probability of F with mean m_0. Let $(Q_n)_{n \in \mathbb{N}}$ be a sequence of ν-pseudo-orthogonal polynomials such that Q_n is of degree $|n|$. Then the following statements are equivalent*:

(i) *The generating function of the (Q_n) is exponential; i.e. there exist a real $r > 0$ and two analytic functions*: b: $B(0, r) \to \mathbb{R}^d$, c: $B(0, r) \to \mathbb{R}$, *such that, for all $z \in B(0, r)$*,

$$\sum_{n \in \mathbb{N}^d} \frac{z^n}{n!} Q_n(x) = \exp\{\langle b(z), x\rangle + c(z)\}.$$

(ii) *There exists $A \in Gl(\mathbb{R}^d)$ such that, for all $n \in \mathbb{N}^d$*,

$$Q_n(x) = Q_0(x)\, P_{A,n}(x) = Q_0(x)\; f_\nu^{(n)}(x, m_0)(Ae_1, ..., Ae_d).$$

In this case $b(z) = \psi_\nu(Az + m_0)$ and $c(z) = -k_\nu(\psi_\nu(Az + m_0))$.

Proof. A detailed proof is in Pommeret (1996). We just give here a shorter version and we start with a lemma.

LEMMA 2.1. *For all $A \in GL(\mathbb{R}^d)$ and for all $(\theta, z) \in \Theta(\nu) \times I$, we have*:

(i) $k_{A^{-1}(\nu)}(\theta) = k_\nu({}^tA^{-1}\theta)$ *and* $\psi_{A^{-1}}(m) = {}^tA\psi_\nu(Am)$,

(ii) $f_\nu^{(n)}(x, m)(Ae_1, ..., Ae_d) = f_{A^{-1}(\nu)}^{(n)}(A^{-1}x, A^{-1}m)(e_1, ..., e_d)$,

where $A^{-1}(\nu)$ is the image measure of ν by A^{-1}.

Proof. Part (i) is a classical exponential family's property (see Letac (1992)). Then from (i) we have that $f_\nu(x, m) = f_{A^{-1}(\nu)}(A^{-1}x, A^{-1}m)$ and part (ii) follows. ∎

By Lemma 1.1 part (ii) and Lemma 2.1 part (ii) we easily see that (ii) implies (i). Conversely we first may show a technical result:

LEMMA 2.2. *With assumptions of Theorem* 2.1, *there exists an open set* $B(0, r)$ *such that* $\forall x,\ y \in B(0, r)$,

$$\int \sum_{n,\,q} \frac{x^n y_q}{n!\,q!} P_n(x)\, P_q(x)\, \nu(dx) = \sum_{n,\,q;\ |n| = |q|} \frac{x^n y^n}{n!\,q!} \int P_n(x)\, P_q(x)\, \nu(dx)$$

Then by Lemma 2.2 we obtain

$$\int \left(\sum_{n \in \mathbb{N}^d} \frac{z^n}{n!} Q_n(x) \right) \nu(dx) = \int Q_0^2(x)\, \nu(dx) = 1.$$

Thus $c(z) = -k_\nu(b(z))$. The (Q_{e_i}) being polynomials with degree one, there exist $(\tilde{A}, v) \in GL(\mathbb{R}^d) \times \mathbb{R}^d$ such that, putting $\mathcal{Q}(x) = {}^t(Q_{e_1}(x), ..., Q_{e_d}(x))$, then $\mathcal{Q}(x) = \tilde{A}x + v$. As previously we have that

$$\int \left(\sum_{n \in \mathbb{N}^d} \frac{z^n}{n!} Q_n(x) \right) \mathcal{Q}(x)\, \nu(dx) = \left(\int {}^t\mathcal{Q}(x)\, \mathcal{Q}(x)\, \nu(dx) \right) z$$
$$= \int \exp\{\langle b(z), x\rangle - k_\nu(b(z))\}\, \mathcal{Q}(x)\, \nu(dx), \tag{8}$$

and comparing (8) with the following two equalities:

$$\int {}^t\mathcal{Q}(x)\, \mathcal{Q}(x)\, \nu(dx) = \tilde{A} V_F(m_0)\, {}^t\tilde{A}$$

$$\int \exp\{\langle t(z), x\rangle - k_\nu(b(z))\}\, \mathcal{Q}(x)\, \nu(dx) = \tilde{A}(k'_\nu(b(z)) - m_0),$$

gives $b(z) = \psi_\nu(V_F(m_0)\, {}^t\tilde{A}z + m_0)$.

Thus $Q_n(x) = f_\nu^{(n)}(x, m_0)(V_F(m_0)^t\, \tilde{A}e_1, ..., V_F(m_0)^t\, \tilde{A}e_d) = P_{A,\,n}(x)$ where $A = V_F(m_0)\, {}^t\tilde{A}$. ▮

Now, as in extension of Feinsilver (1986), we characterize the orthogonality and the pseudo-orthogonality of the polynomials $P_{A,\,n}$, after the following definition: a *NEF* on $\mathbb{R}^d$ $(d > 1)$ will be said *reducible* if there exist two measures μ_1 and μ_2 respectively on $\mathbb{R}^q$ and $\mathbb{R}^{d-q}$ such that $F = F(\mu_1 \otimes \mu_2) = F(\mu_1) \otimes F(\mu_2)$. Obviously, a *NEF* is *irreducible* if it is not reducible. We have:

THEOREM 2.2. *Let F be a NEF on $\mathbb{R}^d$, ν a probability in F with mean m_0 and let A be in $GL(\mathbb{R}^d)$. Then we have the following statements*:

(i) *The $(P_{A,n})_{n\in\mathbb{N}^d}$ are ν-pseudo-orthogonal if and only if F is quadratic*

(ii) *If F is simple quadratic and if $A^{-1}V_F(m_0)\,{}^tA^{-1}$ is diagonal then the $(P_{A,n})_{n\in\mathbb{N}^d}$ are ν-orthogonal. The conversely occurs when F is irreducible.*

Proof. We may suppose $A=id$ (the identity), Lemma 2.1 giving the general case. We content ourselves only with showing part (i) since part (ii) follows this proof (see Pommeret 1996). Let $P_n=P_{id,n}$. By Lemma 1.1 part (ii) and Lemma 2.2 we obtain $\forall m, z\in B(m_0, r)$,

$$\exp\{k_\nu(\psi_\nu(m)+\psi_\nu(z))-k_\nu(\psi_\nu(m))-k_\nu(\psi_\nu(z))\}$$
$$=\sum_{n,q;\,|n|=|q|}\frac{(m-m_0)^n\,(z-m_0)^q}{n!q!}\int P_n(x)\,P_q(x)\,\nu(dx). \tag{9}$$

Differentiating two times (9) with respect to m and taking $m=m_0$ yields F quadratic.

Conversely by Lemma 1.1 part (ii) we have that

$$\exp\{\langle\theta, x\rangle\}=\sum_{n\in\mathbb{N}^d}\frac{(k'_\nu(\theta)-m_0)^n}{n!}P_n(x)\exp\{k_\nu(\theta)\}.$$

By differentiating this equality with respect to θ and by identification we obtain

$$x_iP_n(x)=\sum_{|n|-1\leqslant|q|\leqslant|n|+1}\alpha_qP_q(x), \tag{10}$$

where $(\alpha_q)_{q\in\mathbb{N}^d}\in\mathbb{R}$. By (10) and by induction on $|n|$ we obtain the following three results:

$$\forall n\in\mathbb{N}^d\backslash\{0\},\int P_n(x)\,\nu(dx)=0;$$

There exist $(\alpha^n_{p,q})\in\mathbb{R}$ such that

$$\forall q/|q|<|n|:x^qP_n(x)=\sum_{p;\,|n|-|q|\leqslant|p|\leqslant|n|+|q|}\alpha^n_{p,q}P_p(x);$$

There exist $(\alpha_q)\in\mathbb{R}$ such that

$$P_n(x)=\alpha_nx^n+\sum_{q;\,|q|<|n|}\alpha_qx^q.$$

Thus $\int P_n(x)\,P_q(x)\,\nu(dx)=0$ as soon as $|n|\neq|q|$. ∎

COROLLARY 2.1 (Pommeret, 1996). *With the assumptions of Theorem 2.2, if F is irreducible the two following statements are equivalent*:

(i) *The polynomials* $(P_{A,n})_{n \in \mathbb{N}^d}$ *are* v*-pseudo-orthogonal and the polynomials* $(P_{A,l})_{l \in \mathbb{N}^d,\, |l| \in \{1,2\}}$ *are* v*-orthogonal.*

(ii) *The polynomials* $(P_{A,n})_{n \in \mathbb{N}^d}$ *are* v*-orthogonal.*

3. DIAGONAL AND PSEUDO-DIAGONAL BHATTACHARYYA MATRICES

Let us first remark that the (pseudo-)orthogonality of the polynomials $L_{A,n}$ corresponds exactly to the (pseudo-)diagonality of the associated Bhattacharyya matrix. We consider the pseudo-diagonality and to simplify the notations we use the following results:

LEMMA 3.1. *The following statements are equivalent*:

(i) *There exists* $A \in GL(\mathbb{R}^d)$, *such that* $J_A(z)$ *is pseeudo-diagonal.*

(ii) *For all* $A \in GL(\mathbb{R}^d)$, $J_A(z)$ *is pseudo-diagonal.*

Proof. Let A and $\tilde{A}$ be in $GL(\mathbb{R}^d)$. From the linearity of the derivation we get for all $(x, z) \in \mathbb{R}^d \times I$,

$$\begin{aligned}
L_{\tilde{A},n}(x,z) &= \frac{h_v^{(n)}(x,z)(\tilde{A}A^{-1}Ae_1, ..., \tilde{A}A^{-1}Ae_d)}{h_v(x,z)} \\
&= \sum_{p \in \mathbb{N}^d;\, |p|=|n|} \alpha_p(A, \tilde{A}) \frac{h_v^{(p)}(x,z)(Ae_1, ..., Ae_d)}{h_v(x,z)} \\
&= \sum_{p \in \mathbb{N}^d;\, |p|=|n|} \alpha_p(A, \tilde{A})\, L_{A,p}(x,z),
\end{aligned}$$

where the $\alpha_p(A, \tilde{A})$ are real numbers. It follows that the pseudo-orthogonality of the $(L_{A,n})$ is equivalent to the pseudo-orthogonality of the $(L_{\tilde{A},n})$. ∎

Now, to simplify, let $L_n(x,z) = L_{id,n}(x,z)$ and $J(z) = J_{id}(z)$. The following theorem generalizes Shanbhag's.

THEOREM 3.1. *The following four statements are equivalent*:

(i) *For all* $z \in I$, $J(z)$ *is pseudo-diagonal.*

(ii) *For all* $z \in I$, $J^{1,2}(z) = 0$ *and* $J^{2,3}(z) = 0$.

(iii) *There exists $z \in I$ such that $J(z)$ is pseudo-diagonal.*

(iv) *F is quadratic and there exists $(U, v) \in GL(\mathbb{R}^d) \times \mathbb{R}^d$ such that, for all $z \in I$, $t(z) = \psi_v(Uz + v)$.*

Proof.

(i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) are obvious.

(ii) $\Rightarrow$ (iv) For all $p \in \mathbb{N}^d$, $\partial^{|p|}/\partial z^p$ denotes the partial derivative with respect to z_1 (p_1 times), ..., z_d (p_d times). We begin with a lemma:

LEMMA 3.2. *With the previous notations, writing*

$$L_n(x, z) = \sum_{\substack{q \in \mathbb{N}^d \\ |q| \leqslant |n|}} c_q(z) x^q,$$

where $(c_q)_{n \in \mathbb{N}^d}$ are real numbers, for all $(n, p) \in (\mathbb{N}^d)^2$ we have that

$$\int L_n(x, z)\, L_p(x, z)\, P(t(z), v)(dx) = \sum_{\substack{q \in \mathbb{N}^d \\ |q| \leqslant |n|}} c_q(z) \frac{\partial^{|p|}}{\partial z^p} \left\{ \int x^q P(t(z), v)(dx) \right\}.$$

Proof of the lemma. We can easily see that $\int L_n(x, z)\, x^q P(t(z), v)(dx) = \partial^{|n|}/\partial z^n \left\{ \int x^q P(t(z), v)(dx) \right\}$, and the result follows from the linearity of the scalar product. ▮

Let $t(z) = {}^t(t_1(z), ..., t_d(z))$ and $\langle x_i \rangle = \int x_i\, P(t(z), v)(dx)$. By Lemma 3.2, we obtain for all $p \in \mathbb{N}^d$

$$\int L_{e_i}(x, z\, L_p(x, z)\, P(t(z), v)(dx) = \sum_{s=1}^{d} \frac{\partial t_s}{\partial z_i}(z) \frac{\partial^{|p|}}{\partial z^p} \langle x_s \rangle.$$

In particular, if $|p| = 2$, $J^{1,2}(z) = 0$ implies that:

$$\sum_{s=1}^{d} \frac{\partial t_s}{\partial z_i}(z) \frac{\partial^2}{\partial z^p} \langle x_s \rangle = 0.$$

Since t' is invertible, then $\langle x_s \rangle = \int x_s\, P(t(z), v)(dx)$ is a polynomial in z of degree 1 and therefore there exist U and v such that

$$\langle x \rangle = k'_v(t(z)) = Uz + v. \tag{11}$$

By derivation, we obtain $U = {}^t t'(z)\, k''_v(t(z))$ and then U belongs to $GL(\mathbb{R}^d)$. By (11) we obtain

$$t(z) = \psi_v(Uz + v). \tag{12}$$

Proceding in a similar manner to previously, by $J^{2,3}=0$ we obtain for all $p\in\mathbb{N}^d$ such that $|p|=3$:

$$\int L_{e_i+e_j}(x,z)\,L_p(x,z)\,P(t(z),v)(dx)=\sum_{s,q=1}^{d}\frac{\partial t_s}{\partial z_i}(z)\,\frac{\partial t_q}{\partial z_i}(z)\,\frac{\partial^3}{\partial z^p}\langle x_s x_q\rangle=0.$$

Then $\langle x\,{}^t x\rangle=(\langle x_s x_q\rangle)_{s,q=1\cdots d}$ is a polynomial matrix in z of degree $\leqslant 2$, and the following equality

$$\langle x\,{}^t x\rangle=V_F(k_v'(t(z)))+\langle x\rangle\,{}^t\langle x\rangle,$$

yields there exists an open set $\subset M_F$ under which $V_F(m)$ has degree $\leqslant 2$. Hence F is quadratic.

(iv) $\Rightarrow$ (i) Let z_0 be in I and let $\mu=P(t(z_0),v)$ and $m_0=Uz_0+v$. Then

$$f_\mu(x,Uz+v)=\frac{f_v(x,Uz+v)}{f_v(x,m_0)}=\frac{h_v(x,z)}{h_v(x,z_0)},$$

and using the definition of the $L_n=L_{id,n}$ we obtain

$$L_n(x,z_0)=f_\mu^{(n)}(x,Uz_0+v)(Ue_1,...,Ue_d)=P_{U,n}(x)$$

and by Theorem 2.2, the polynomials $(L_n(x,z_0))$ are μ-pseudo-orthogonal, i.e. $J(z_0)$ is pseudo-diagonal. Since z_0 is arbitrary we obtain (i).

It remains to prove (iii) $\Rightarrow$ (iv). Let z_0 be in I such that $J(z_0)$ is pseudo-diagonal. Denote $\mu=P(t(z_0),v)$ and $\tilde{t}(z)=t(z)-t(z_0)$. Then by Lemma 1.1

$$\sum_{n\in\mathbb{N}^d}\frac{(z-z_0)^n}{n!}L_n(x,z_0)=\exp\{\langle t(z)-t(z_0),x\rangle-k_v(t(z))+k_v(t(z_0))\}$$
$$=\exp\{\langle\tilde{t}(z),x\rangle-k_\mu(\tilde{t}(z))\}.$$

Thus, the polynomials $(L_n(x,z_0))_{n\in\mathbb{N}^d}$ are μ-pseudo-orthogonal with exponential genereating function. By Theorem 2.1, there exists $A\in GL(\mathbb{R}^d)$ such that

$$L_n(x,z_0)=P_{A,n}(x)=f_\mu^{(n)}(x,m_0)(Ae_1,...,Ae_d)$$

with $m_0=k_\mu'(0)=k_v'(t(z_0))$.

Hence $\tilde{t}(z+z_0)=\psi_\mu(Az+m_0)=\psi_v(Az+m_0)-\psi_v(m_0)$ and consequently $t(z)=\psi_v(Az+m_0-Az_0)=\psi_v(Uz+v)$ where $U=A$ and $v=m_0-Az_0$. Furthermore, the $(P_{A,n})$ being pseudo-orthogonal, F is quadratic. ∎

Now we take $A\in GL(\mathbb{R}^d)$, not necessarily the identity, and we give a characterization of the diagonality of the Bhattacharyya matrices.

THEOREM 3.2. *If f is irreducible, then the following three statements are equivalent*:

(i) $J_A(z_0)$ *is diagonal.*

(ii) *For all* $z \in I$, $J_{A^{1,2}}(z)=0$ *and* $J_{A^{2,3}}(z)=0$ *and* $J_{A^{1,1}}(z_0)$ *and* $J_{A^{2,2}}(z_0)$ *are diagonal.*

(iii) *F is simple quadratic and there exists* $(U, v) \in GL(\mathbb{R}^d) \times \mathbb{R}^d$ *such that, for all* $z \in I$, $t(z)=\psi_v(Uz+v)$ *and* $(UA)^{-1} V_F(Uz_0+v)\, {}^t(UA)^{-1}$ *is diagonal.*

Proof.

(i) $\Rightarrow$ (ii) It clearly comes from Theorem 3.1.

(ii) $\Rightarrow$ (iii) By Theorem 3.1 F is quadratic and there exists $(U, v) \in GL(\mathbb{R}^d) \times \mathbb{R}^d$ such that

$$t(z)=\psi_v(Uz+v). \tag{13}$$

For $z_0 \in I$, let $\mu = P(t(z_0), v)$. Then by Lemma 1.1,

$$\begin{aligned}\sum_{n \in \mathbb{N}^d} \frac{(z-A^{-1}z_0)^n}{n!} L_{A,n}(x, z_0) &= \exp\{\langle t(Az)-t(z_0), x\rangle - k_v(t(Az))+k_v(t(z_0))\} \\ &= \exp\{\langle \tilde{t}(Az), x\rangle - k_\mu(\tilde{t}(Az))\},\end{aligned}$$

where $\tilde{t}(z)=t(z)-t(z_0)$.

Since the $L_{A,n}(z_0)$ are μ-pseudo-orthogonal and have an exponential generating function, then by Theorem 2.1, there exists $\tilde{A} \in GL(\mathbb{R}^d)$ such that:

$$L_{A,n}(x, z_0) = f_\mu^{(n)}(x, m_0)(\tilde{A}e_1, ..., \tilde{A}e_d),$$

where $m_0 = Uz_0+v$ and $\tilde{t}(A(z+z_0)) = \psi_\mu(\tilde{A}z+m_0) = \psi_v(\tilde{A}z+m_0)-\psi_v(m_0)$. Hence,

$$t(z)=\psi_v(\tilde{A}A^{-1}z-Az_0+m_0). \tag{14}$$

Since $J_A^{1,1}(z_0)$ and $J_A^{2,2}(z_0)$ are diagonal the $(d^2+3d)/2$ polynomials $(L_{A,n}(x, z_0))_{|n|=1,2}$ are μ-orthogonal. Then by Corollary 2.1 the polynomials $(L_{A,n}(x, z_0))_{n \in N^d}$ are μ-orthogonal and by Theorem 2.2 F is simple quadratic and $\tilde{A}^{-1} V_F(Uz_0+v)\, {}^t\tilde{A}^{-1}$ is diagonal. By (13) and (14) we have $\tilde{A}=UA$ and it follows the diagonality of $(UA)^{-1} V_F(m_0)\, {}^t(UA)^{-1}$.

(iii) $\Rightarrow$ (i) By Theorem 2.2, F quadratic implies that the polynomials $(L_{A,n}(z_0))_{n \in \mathbb{N}^d}$ are μ-pseudo-orthogonal with $\mu = P(t(z_0), v)$. As above, there exists $\tilde{A} = UA$ such that $L_{A,n}(z_0) = P_{\tilde{A},n}(x)$. F being simple quadratic and $\tilde{A}^{-1} V_F(Uz_0 + v)^t \tilde{A}^{-1}$ being diagonal, by Theorem 2.2 the polynomials $(P_{\tilde{A},n})_{n \in \mathbb{N}^d}$ are μ-orthogonal, i.e. $J_A(z_0)$ is diagonal. ∎

4. EXAMPLES OF MULTIDIMENSIONAL DIAGONAL BHATTACHARYYA MATRICES

Restricting ourself to the class of simple quadratic *NEF*, we consider here the diagonal Bhattacharyya matrices on $\mathbb{R}^d$. To simplify the notations we assume that $t(z) = \psi_v(z)$ and $I = M_F$. Clearly we have $h_v(x, m) = f_v(x, m)$. Since $V_F(m)$ is symmetric and positive, for all $m \in M_F$ there exists $A \in GL(\mathbb{R}^d)$ such that $A^{-1} V_F(m)\ {}^t A^{-1}$ is diagonal. In this case, we have the diagonal Bhattacharyya matrix:

$$J_A(m) = \begin{pmatrix} \|L_{A,e_1}\|^2 & & 0 & \cdots & 0 \\ & \|L_{A,e_2}\|^2 & & & \vdots \\ 0 & & \ddots & & 0 \\ \vdots & & & \|L_{A,n}\|^2 & \\ 0 & \cdots & 0 & & \ddots \end{pmatrix}$$

where $\|\cdot\|$ denotes the $L^2(P(t(m), v))$ norm; i.e. $\|L_{A,n}\|^2 = \int (L_{A,n}(x, m))^2 P(t(m), v)(dx)$. Thus, in the simple quadratic case, investigation of $J_A(m)$ reduces simply to investigation of the $L_{A,n}$'s norm.

We recall briefly here the $2d+4$ types of simple quadratic *NEF* on $\mathbb{R}^d$ described by Casalis (1996):

The $d+1$ Poisson Gaussian types. For $k = 0, ..., d$, $V_F(m) = diag(m_1, ..., m_k, 1, ..., 1)$ on $M_F =]0; +\infty[^k \times \mathbb{R}^{d-k}$, where *diag* denotes a diagonal matrix. There are composed by the distributions of $(X_1, ..., X_d)$, where $X_1, ..., X_d$ are independents. $X_1, ..., X_k$ have a Poisson distribution and $X_{k+1}, ..., X_d$ are Gaussian variables with variance 1.

The $d+1$ negative multinomial gamma types. For $k = 0, ..., d$, $V_F(m) = m\ {}^t m + diag(m_1, ..., m_k, 0, m_{k+1}, ..., m_{k+1})$ on $M_F =]0; +\infty[^{k+1} \times \mathbb{R}^{d-k-1}$. There are composed by the distributions of $(X_1, ..., X_d)$, where $(X_1, ..., X_k)$ has a negative multinomial distribution with shape parameter 1. X_{k+1} given $(X_1, ..., X_k)$ is gamma distributed with shape parameter $\sum_{i=1}^k X_i + 1$ and $(X_{k+2}, ..., X_d)$ given $(X_1, ..., X_{k+1})$ are $(d-k-1)$ real independent Paussian variables with variance X_{k+1}.

The multinomial type. $V_F(m) = -m\ {}^t m + diag(m_1, ..., m_d)$, $M_F = \{m \in \mathbb{R}^d;\ m_j > 0, \sum_{j=1}^d m_j < 1\}$.

The hyperbolic type. $V_F(m) = m\,{}^t m + diag(m_1, ..., m_{d-1}, \sum_{i=1}^{d-1} m_i + 1)$ on $M_F =]0; +\infty[^{d+1} \times \mathbb{R}$. F is composed by the distributions of $(X_1, ..., X_d)$ where $(X_1, ..., X_{d-1})$ has a negative-multinomial distrubtion and X_d given $(X_1, ..., X_{d-1})$ has the hyperbolic cosine distribution with power convolution parameter $\sum_{i=1}^{d-1} x_i + 1$.

Each type is composed by one *NEF* of the same name and its affinities and convolution powers; i.e. $F(v)$ and $F(\mu)$ are said to be of the same type if there exist an affinity $\phi: x \mapsto Ax + b$ and a positive real number λ such that $v = \phi(\mu^{*\lambda})$ (where $*$ denotes the convolution product).

The following proposition gives the $L_{A,n}$'s norm and generalizes a similar result established by Seth (1949) on $\mathbb{R}$.

PROPOSITION 4.1. *Let $F = F(v)$ be a simple quadratic NEF on $\mathbb{R}^d$ such that $V_F(m) = am\,{}^t m + B(m) + C$. Then for all $(m, A) \in M_F \times GL(\mathbb{R}^d)$ such that $A^{-1} V_F(m)\,{}^t A^{-1} = diag(h_1, ..., h_d)$, we have*:

$$\|L_{A,n}\|^2 = \begin{cases} n! \left(\dfrac{1}{h_1}\right)^{n_1} \cdots \left(\dfrac{1}{h_d}\right)^{n_d}, \\ \qquad \textit{if} \quad F \in \textit{Poisson Gaussian type}. \\ n!\,a^{|n|}\,\dfrac{\Gamma(1/a + |n|)}{\Gamma(1/a)} \left(\dfrac{1}{h_1}\right)^{n_1} \cdots \left(\dfrac{1}{h_d}\right)^{n_d}, \\ \qquad \textit{if} \quad F \in \textit{negative multidimensional gamma or hyperbolic type}. \\ n!\,|a|^{|n|}\,\dfrac{\Gamma(-(1/a)+1)}{\Gamma(-(1/a) - |n| + 1)} \left(\dfrac{1}{h_1}\right)^{n_1} \cdots \left(\dfrac{1}{h_d}\right)^{n_d} 1I_{|n| \leqslant (-1/a)}, \\ \qquad \textit{if} \quad F \in \textit{multinomial type}. \end{cases}$$

where

$$\frac{\Gamma(1/a + |n|)}{\Gamma(1/a)} = \left(\frac{1}{a} + |n| - 1\right) \cdots \left(\frac{1}{a}\right).$$

Proof. We summarize the proof in three steps:

(i) writing $Q_n(x, z) = h^{(n)}_{A^{-1}(v)}(x, z)(e_1, ..., e_d) / h_{A^{-1}(v)}(x, z)$ and using Lemma 3.1 we have that

$$\int_{\mathbb{R}^d} (L_{A,n}(x, m))^2\, h_v(x, m)\, v(dx)$$

$$= \int_{\mathbb{R}^d} (Q_n(x, A^{-1}m))^2\, h_{A^{-1}(v)}(x, A^{-1}(v)(dx),$$

and the Q_n are $P(t(A^{-1}m), A^{-1}(v))$-orthogonal.

(ii) by the following two equalities,

$$Q_{e_i}(x, m) = (V_{A^{-1}(F)}^{-1}(m)(x-m))_i, \tag{15}$$

$$\frac{\partial}{\partial m_i}(Q_n(x, m)) = Q_{n+e_i}(x, m) - Q_n(x, m)\, Q_{e_i}(x, m), \tag{16}$$

we may show by induction on $|n|$ that:

$$\begin{aligned} &Q_{e_i}(x, A^{-1}m)\, Q_n(x, A^{-1}m) \\ &\quad = \left(\frac{1}{h_i}\right)\Bigg(Q_{n+e_i}(x, A^{-1}m) + \sum_{j=1}^{d} n_j\, Q_{n-e_j}(x, A^{-1}m) \\ &\qquad + a n_i(|n|-1)\, Q_{n-e_i}(x, A^{-1}m) \\ &\qquad + \sum_{j,s}^{d} n_j \frac{\partial}{\partial m_j}(V_{A^{-1}F}(A^{-1}m))_{is}\, Q_{n+e_s-e_j}(x, A^{-1}m)\Bigg). \end{aligned} \tag{17}$$

(iii) By (16) we obtain

$$\begin{aligned} &\int_{\mathbb{R}^d} (Q_n(x, A^{-1}m))^2\, h_{A^{-1}(v)}(x, A^{-1}m)\, A^{-1}(v)(dx) \\ &\quad = \int_{\mathbb{R}^d} Q_{n-e_i}(x, A^{-1}m))\, Q_{e_i}(x, A^{-1}m)\, Q_n(x, A^{-1}m)) \\ &\qquad \times h_{A^{-1}(v)}(x, A^{-1}m)\, A^{-1}(v)(dx), \end{aligned}$$

and the result follows from a last induction, using (17) and the orthogonality of the Q_n. ▮

Remark. The quadratic case, for example the Wishart distributions, giving pseudo-diagonal Bhattacharyya matrices is not approached here. In fact the norm of the pseudo-orthogonal polynomials $L_{A,n}$ is not yet determined.

ACKNOWLEDGMENTS

Thanks are due to professors Casalis and Letac for their helpful comments and suggestions.

REFERENCES

[1] Bhattacharyya, A. (1946). On some analogues of the amount of information and their use in statistical estimation. *Sankhyā* **8**.

[2] Casalis, M. (1996). The $2d+4$ simple quadratic natural exponential families on $\mathbb{R}^d$. *Ann. Statist.*, to appear.

[3] Feinsilver, P. (1986). Some classes of orthogonal polynomials associated with martingales. *Proc. A.M.S.* **98** 298–302.

[4] Letac, G. (1992). *Lectures on natural exponential families and their variance functions.* Instituto de matemática pura e aplicada: Monopgrafias de matemática, **50** Rìo.

[5] Meixner, J. (1934). Orthogonal Polynomsysteme mit einer besonderen Gestalt der erzengenden Function. *J. London Math. Soc.* **9** 6–13.

[6] Morris, C. N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10** 65–82.

[7] Pommeret, D. (1996). Orthogonal polynomials and natural exponential families. *Test* **5** 77–111.

[8] Seth, G. R. (1949). On the variance of estimates. *Ann. Math. Statist.* **20** 1–27.

[9] Shanbhag, D. N. (1972). Some characterizations based on the Bhattacharyya matrix. *J. Appl. Prob.* **9** 580–587.

[10] Shanbhag, D. N. (1979). Diagonality of the Bhattacharyya matrix as a characterization. *Theor. Prob. and Appl.* **24** 430–433.