Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Dualistic structures in information geometry

by

Léonard Todjihounde

Preprint no.: 32 2022

Dualistic structures in information geometry

Leonard TODJIHOUNDE*

1

Abstract

We revisit some important basics of dualistic structures on Riemannian manifolds as a fundamental mathematical concept connecting information geometry, affine geometry and Hessian geometry. Since several statistical manifolds can be seen as warped product of spaces, we conclude this review by some results on warped products of dualistic structures.

1 Introduction

Information geometry provides a geometric approach to families of statiscal models. The key geometric structures are the Fisher metric, the divergence functions and the Amari-Chentov tensor. Dualistic structures are a fundamental mathematical concept of information geometry, specially in the investigation of the natural differential geometric structure of probability distributions families. They were introduced by S. Amari and have been studied by several authors (see [1], [2], [3], [8] and the references therein). Dual flatness constitutes a fundamental mathematical concept connecting information geometry, affine geometry and Hessian geometry. Moreover the existence of dually flat structures points out some topological and geometrical properties of the underlying manifolds (see [4]).
For example compact Riemannian manifolds with trivial or finite fundamental groups do never admit dually flat structures.
Also if a manifold M admits a dually flat structure (g, ∇, ∇^*) and if one of the dual connection, say ∇, is complete, then only the first homotopy group of M is non trivial, and any two points in M can be joined by a ∇-geodesic.
Warped dualistic structures are also of fundamental importance in information geometry since several metrics on statistical models are modelled as warped products such as the Fisher metric on the multivariate Gaussian distributions space (the so-called Takano Gaussian space), the denormalizations of the Fisher metric, etc.
It has been showed in [10] that the set of multivariate Gaussian distributions

[1]Permanent address: Institut de Mathématiques et de Sciences Physiques (IMSP), Porto-Novo, Benin. Email: leonardt@imsp-uac.org , leonardt67@gmail.com
The author thanks warmly the Max-Planck Institute for Mathematics in the Sciences for his financial and logistical support during the stay where this work has been done.

on $\mathbb{R}^n$ with mean zero, equipped with the L^2-Wasserstein metric, has a cone structure. Also interesting relations have been obtained in [9] between Fisher metrics of location scale models and warped product metrics.

2 Dual connections on Riemannian manifolds

Let (M, g) be a Riemannian manifold and ∇ an affine connection on M.
A connection ∇^* is called conjugate connection (or dual connection) of ∇ with respect to the metric g if, for arbitrary $X, Y, Z \in \chi(M)$,

$$X \cdot g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z) \quad (*)$$

The duality relation between affine connections on a given Riemannian manifold (M, g) is symmetric; that is

$$(\nabla^*)^* = \nabla \; .$$

The triple (g, ∇, ∇^*) satisfying $(*)$ is then called a dualistic structure on M and (M, g, ∇, ∇^*) is called a statistical manifold.

Let (g, ∇, ∇^*) be a dualistic structure on M. Then $\nabla^{(0)} = \frac{1}{2}(\nabla + \nabla^*)$ is a metric connection on (M, g) anf if moreover ∇ and ∇^* are torsion-free then $\nabla^{(0)}$ is the Levi-Civita connection of (M, g).

Let $\gamma : t \longmapsto \gamma(t)$ be a curve in M with between the points $p, q \in M$. Denote by Π_γ and Π^*_γ the parallel translation along the curve γ w.r.t. ∇ and its dual ∇^* respectively. It holds:

$$g(\Pi_\gamma(X), \Pi^*_\gamma(Y))_{!q} = g(X, Y)_{!p} \; , \; \forall \; X, Y \in \chi(M) \; .$$

Indeed if we denote by $X(t)$ the ∇-parallel translation of X at the point $\gamma(t)$, and by $Y(t)$ the ∇^*-parallel translation of Y at the point $\gamma(t)$, we get

$$\begin{aligned} \frac{d}{dt} g(X(t), Y(t)) &= \frac{\partial}{\partial t} \cdot g(X(t), Y(t)) \\ &= g(\nabla_{\partial t} X(t), Y(t)) + g(X(t), \nabla^*_{\partial t} Y(t)) \\ &= 0 + 0 = 0 \; , \end{aligned}$$

since the vector fields X and Y are respectively ∇-parallel and ∇^*-parallel along γ. Thus $g(X(t), Y(t))$ is a constant function. $\square$

The previous equality can be interpreted as a generalization of the invariance of the inner product under parallel translation through metric connections.

Conjugate connections arise from affine differential geometry and from geometric theory of statistical inferences and it is a natural generalization of Levi-Civita connections.

The torsion tensors T^{∇} and T^{∇^*} of ∇ and ∇^*, respectively, satisfy the relation:

$$g(T^{\nabla}(X,Y),Z) = g(T^{\nabla^*}(X,Y),Z) + (\nabla^* g)(X,Y,Z) - (\nabla^* g)(Y,X,Z)$$

for $X,Y,Z \in \chi(M)$. IT follows then

$$T^{\nabla} = T^{\nabla^*} \text{ if and only if } \nabla^* g(X,Y,Z) = \nabla^* g(Y,X,Z) \; .$$

In particular: If $\nabla^* g = 0$, then $T^{\nabla} = T^{\nabla^*}$.

Let (g, ∇, ∇^*) be a dualistic structure on M, and let R and R^* be the curvature tensors of ∇ and ∇^* respectively. We have:

$$g(R(X,Y)Z, W) = -g(R^*(X,Y)W, Z) \; , \; \forall \; X,Y,Z,W \in \chi(M).$$

Therefore

$$R = 0 \text{ if and only if } \; R^* = 0 \; .$$

The manifold M endowed with a dualistic structure (g, ∇, ∇^*) is called a *dually flat space* if both dual connections ∇ and ∇^* are torsion free and flat; that is the curvature tensors with respect to ∇ and ∇^* respectively vanishe identically.

Remark 2.1:
The flatness of a statistical manifold does not imply that the manifold is Euclidean in the geometrical sense, because the Riemannian curvature due to the Levi-Civita connections does not necessarily vanish.

Flat Dualistic structures connect information geometry, affine differential geometry and Hessian geometry in the following sense:

Let ∇ and ∇^* be dually flat and torsion-free connections on an n-dimensional Riemannian manifold (M,g), $(u^1, \cdots, u^n)$ an affine coordinates system w.r.t. ∇.
The vector fields $\partial_i := \frac{\partial}{\partial u^i}$ are then parallel w.r.t. ∇.
We define vector fields ∂^j via

$$g(\partial_i, \partial^j) = \delta_i^j \; .$$

For any vector field X,we have:

$$0 = X.g(\partial_i, \partial^j) = g(\nabla_X \partial_i, \partial^j) + g(\partial_i, \nabla^*_X \partial^j)$$

and since ∂_i is parallel w.r.t. ∇, we get that ∂^j is parallel w.r.t. ∇^*.
Since ∇^* is torsion-free, then also $[\partial^j, \partial^k] = 0$ for all j and k, and thus we may

find ∇^*-affine coordinates $(v_1, \cdots, v_n)$ with $\partial^j = \frac{\partial}{\partial v_j}$.

The metric tensor g in terms of u-coordinates and v-coordinates is given by:

$$g_{ij} = g(\partial_i, \partial_j) \text{ and } g^{ij} = g(\partial^i, \partial^j)$$

From the duality relation $g(\partial_i, \partial^j) = \delta_i^j$ it follows

$$\frac{\partial v_j}{\partial u^i} = g_{ij} \ , \ \frac{\partial u^i}{\partial v_j} = g^{ij} \ .$$

The transition rules between u-coordinates and v-coordinates are given by:

$$\partial^j = (\partial^j u^i)\partial_i = g^{ij}\partial_i \ , \ \partial_i = (\partial_i v_j)\partial^j = g_{ij}\partial^j \ .$$

Proposition 2.1
There exist strictly convex potential functions $\phi(v)$ and $\psi(u)$ satisfying

$$v_i = \partial_i \psi(u) \ , \ u^i = \partial^i \phi(v) \ ,$$

as well as

$$g_{ij} = \partial_i \partial_j \psi \ , \ g^{ij} = \partial^i \partial^j \phi \ .$$

Proof:
The equation $v_i = \partial_i \psi(u)$ can be solved iff

$$\partial_i v_j = \partial_j v_i \ ,$$

which holds from the symmetry $g_{ij} = g_{ji}$. Moreover $g_{ij} = \partial_i \partial_j \psi$, thus ψ is strictly convex.

The function ϕ can be found by the argument or by duality; in fact, we can simply define ϕ by

$$\phi(v) := u^i v_i - \psi(u) \ ,$$

and it holds

$$\partial^i \phi = u^i + \frac{\partial u^j}{\partial v_i} v_j - \frac{\partial u^j}{\partial v_i} \frac{\partial}{\partial u^j} \psi = u^i \ .$$

□

Since ψ and ϕ are strictly convex, the relation

$$\phi(v) + \psi(u) = u^i v_i$$

means that they are related by Legendre transformations,

$$\phi(v) = \max_u (u^i v_i - \psi(u)) \ , \ \psi(u) = \max_v (u^i v_i - \phi(v)) \ .$$

A statistical structure on a manifold M is equivalent to a pair $(g, \mathcal{T})$, where is a Riemannian metric on M and $\mathcal{T}$ a 3-tensor that is symmetric in all three arguments.
Indeed let ∇ and ∇^* be dual and torsion-free connections on a Riemannian manifold (M, g). Then the 3-tensor $\mathcal{T}$ defined by

$$\mathcal{T} = \nabla - \nabla^*$$

is a symmetric 3-tensor on (M, g). We have

$$\mathcal{T}(X, Y, Z) = g(\nabla_X Y - \nabla^*_X Y, Z) = 2g(\nabla^{(0)}_X Y, Z) - 2g(\nabla_X Y, Z) ,$$

where $\nabla^{(0)}$ is the Levi-Civita connection of (M, g).

By putting the associated 2-tensor $\mathbf{T}$ by $\mathcal{T}(X, Y, Z) = g(\mathbf{T}(X, Y), Z)$, we have

$$g(\mathbf{T}(X, Y), Z) = \nabla_X g(Y, Z) .$$

Conversely a statistical structure $(g, \mathcal{T})$ yields a torsion-free dualistic structure on M.
Indeed ∇ and $\nabla*$ defined by

$$\nabla_X Y = \nabla^{(0)}_X Y - \frac{1}{2}\mathbf{T}(X, Y) \text{ and } \nabla^*_X Y = \nabla^{(0)}_X Y + \frac{1}{2}\mathbf{T}(X, Y) ,$$

are torsion-free connections on M that are dual to each other w.r.t. the Riemannian metric g.

More generally, from such a pair $(g, \mathcal{T})$ we can define a entire family of torsion-free connections by puting, for $-1 \leq \alpha \leq 1$,

$$\nabla^{(\alpha)}_X Y = \nabla^{(0)}_X Y - \frac{\alpha}{2}\mathbf{T}(X, Y) ;$$

or locally

$$\Gamma^{(\alpha)}_{ijk} = \Gamma^{(0)}_{ijk} - \frac{\alpha}{2}\mathbf{T}_{ijk} ,$$

with

$$T_{ijk} = \mathbb{E}(\partial_i l \partial_j l \partial_k l) = \partial_i \partial_j \partial_k \psi ,$$

where ψ is the potential function associated to the affine coordinates system.

The connections $\nabla^{(\alpha)}$ and $\nabla^{(-\alpha)}$ are the dual to each other w.r.t. g.

The torsion-free connection $\nabla^{(1)}$ is flat on the families of exponential distributions with their Fischer metrics; it is called exponential connection, e-connection for short. Its dual connection $\nabla^{(-1)}$ is called mixture-connection, m-connection for short.

3 Dualistic structures and conformal change of metrics

Let (M, g) be a Riemannian manifold. Homothetic changes of the metric g preserve the dualistic structure. More precisely, if (M, g, ∇, ∇^*) is a statistical manifold and $\tilde{g} = cg$, with $c \in \mathbb{R}_{>0}$, then $(M, \tilde{g}, \nabla, \nabla^*)$ is also a statistical manifold. In the general case of conformal changes of the metric g we have the following result:

Proposition 3.1
A conformal change of the Riemannian metric g preserves a dualistic structure on (M, g) if and only if it preserves the Levi-Civita connection of g.
More precisely if (M, g, ∇, ∇^*) is a statistical manifold and $\tilde{g}$ is a conformal metric to g, then $(M, \tilde{g}, \nabla, \nabla^*)$ is a statistical manifold if and only if $\nabla^{(0)} = \tilde{\nabla}^{(0)}$, where $\nabla^{(0)}$ and $\tilde{\nabla}^{(0)}$ are the Levi-Civita connections of g and $\tilde{g}$ respectively.

Proof
Let (M, g, ∇, ∇^*) a statistical manifold and $\tilde{g} = \mathrm{e}^{\phi}\, g$, with $\phi \in C^{\infty}(M)$, a conformal metric to g. Then we have

$$\nabla^{(0)} = \frac{1}{2}(\nabla + \nabla^*)\ ,\ \text{where } \nabla^{(0)} \text{ is the Levi-Civita connection of } g\ .$$

The Levi-Civita connection $\tilde{\nabla}^{(0)}$ of $\tilde{g}$ is related to $\nabla^{(0)}$ by:

$$\tilde{\nabla}^{(0)}_X Y = \nabla^{(0)}_X Y + X(\phi)Y + Y(\phi)X - g(X, Y)\,\mathrm{grad}\,\phi\ ,\ \forall\, X, Y \in \chi(M)\ .$$

Thus it holds:

$$\tilde{\nabla}^{(0)} = \frac{1}{2}(\nabla + \nabla^*) + \mathcal{K}_{\phi}\ ,$$

where $\mathcal{K}_{\phi}$ is the $(1, 2)$-tensor defined by

$$\mathcal{K}_{\phi}(X, Y) = X(\phi)Y + Y(\phi)X - g(X, Y)\,\mathrm{grad}\,\phi\ ,\ \forall\, X, Y \in \chi(M)\ .$$

Therefore the connection ∇ and ∇^* are dual with respect to $\tilde{g}$ if and only if $\mathcal{K}_{\phi}(X, Y) = 0\ ,\ \forall\, X, Y \in \chi(M)$. $\square$

4 The Fisher metric on statistical models

Let $\mathcal{S} = \{p(x; \zeta)\ /\ x \in \Omega\ ,\ \zeta \in \mathbb{R}^n\}$ be a set of probability distributions $p(x; \zeta)$ defined on a measurable space Ω and parametrized by $\zeta \in \mathbb{R}^n$. The set $\mathcal{S}$ can be modelled as an n-dimensional manifold by identifying an element $p(x; \zeta)$ of $\mathcal{S}$ with the parameter $\zeta \in E$.

The Fisher metric g on $\mathcal{S}$ is defined by

$$g_{ij}(\zeta) = \mathbb{E}[\partial_i l(x, \zeta) \partial_j l(x, \zeta)] = \int_{\Omega} p(x; \zeta) \partial_i l(x, \zeta) \partial_j l(x, \zeta)\, dx\ ,$$

with $l(x,\zeta) = \log p(x,\zeta)$ and $\partial_k = \frac{\partial}{\partial \zeta^k}$.

The matrix (g_{ij}) is symmetric; that is: $g_{ij} = g_{ji}$ and is positive semidefinite. We make it positive definite by assuming that the vectors ∂_i , $i = 1, \cdots, n$, are linearly independent. In this case g is a Riemannian metric on $\mathcal{S}$. We have also:

$$g_{ij}(\zeta) = -\mathbb{E}[\partial_i \partial_j l(x,\zeta)] = -\int_{\Omega} p(x;\zeta)\partial_i\partial_j l(x,\zeta)\, dx \ .$$

Examples 4.1
1. Let us consider multivariate Gaussian distributions

$$p(x,\zeta) = \frac{1}{\sqrt{2\pi}\sigma}\Pi_{i=1}^{n} \exp\{-\frac{(x_i - m_i)^2}{2\sigma^2}\} \ ,$$

where $\zeta = (\sigma, m_1, \cdots, m_n) \in \mathbb{R}_{>0} \times \mathbb{R}^n$. The set $\mathcal{M} = \{p(x,\zeta) \ / \ x \in \mathbb{R}^n \ ; \ \zeta \in \mathbb{R}_{>0} \times \mathbb{R}^n\}$ could be modelled as an $(n+1)$-dimensional manifold, called Takano Gaussian space.
The components of the Fisher metric G on the Takano space are given by

$$G_{\sigma\sigma} = \frac{2n}{\sigma^2} \ , \ G_{\sigma i} = G_{i\sigma} = 0 \text{ and } G_{ij} = \frac{1}{\sigma^2}\delta_{ij} \ , \ i,j = 1, \cdots, n \ .$$

Equvalently the Fisher metric G on the Takano Gaussian space is given by

$$ds^2 = \frac{1}{\sigma^2} 2n d\sigma_2 + dm_1^2 + \cdots + dm_n^2 \ .$$

The components of α-connections on $(\mathcal{M}, G)$ are given by

$$\Gamma^{(\alpha)}_{ij,k} = 0 \ , \ \Gamma^{(\alpha)}_{ij,\sigma} = \frac{1-\alpha}{\sigma^3}\delta_{ij} \ , \ \Gamma^{(\alpha)}_{ij,k} = -\frac{1+\alpha}{\sigma^3}\delta_{ij} \ ,$$

$$\Gamma^{(\alpha)}_{i\sigma,\sigma} = \Gamma^{(\alpha)}_{\sigma\sigma,i} = 0 \ , \ \Gamma^{(\alpha)}_{\sigma\sigma,\sigma} = -(1+2\alpha)\frac{2n}{\sigma^3} \ .$$

The corresponding α-connections on $(\mathcal{M}, G)$ are then:

$$\nabla^{(\alpha)}_{\partial_i}\partial_j = \delta_{ij}\partial_\sigma \ , \ \nabla^{(\alpha)}_{\partial_i}\partial_\sigma = \nabla^{(\alpha)}_{\partial_\sigma}\partial_i = -\frac{1+\alpha}{\sigma}\partial_i \ , \ \nabla^{(\alpha)}_{\partial_\sigma}\partial_\sigma = -\frac{1+2\alpha}{\sigma}\partial_\sigma \ .$$

Proposition 4.1
$(\mathcal{M}, G, \nabla^{(\alpha)})$ is a space of constant curvature $-\frac{(1-\alpha)(1+\alpha)}{2n}$.

It follows that $(G, \nabla^{(1)}), \nabla^{(-1)})$ is a dually flat struture on the Takano Gaussian space $\mathcal{M}$.

2. **The Gamma manifold**

Consider the family of gamma densities

$$f(x;\mu;\beta) = \frac{(\frac{\beta}{\mu})^{\beta} x^{\beta-1}}{\Gamma(\beta)} \exp(-\frac{x\beta}{\mu}) \ , \ \mu > 0 \ , \ \beta > 0$$

w.r.t. to Lebesgue measure on $\mathbb{R}_{+}$.

The set $\mathcal{G}$ of Gamma densities could be modelled as a 2-dimensional manifold called Gamma manifold; more precisely

$$\mathcal{G} \equiv \{(\mu,\beta) \in \mathbb{R}^2 \ / \ \mu > 0 \ , \ \beta > 0\} = \mathbb{R}_{>0} \times \mathbb{R}_{>0} \ .$$

The Fisher metric g is given in the (μ,β)-parametrization by

$$ds^2 = \frac{\beta}{\mu^2} \, d\mu^2 + \phi(\beta) \, d\beta^2$$

where $\phi(\beta) = \psi'(\beta) - \frac{1}{\beta}$ with $\psi(x) = \frac{\Gamma(x)}{\Gamma'(x)} \approx \ln x - \frac{1}{2x}$.

The components of the Levi-Civita connection are:

$$\Gamma_{111} = -\frac{\beta}{\mu^3} \ , \ \Gamma_{112} = -\frac{1}{2\mu^2} \ , \ \Gamma_{121} = \Gamma_{211} = \frac{1}{\mu^3} \ ,$$

$$\Gamma_{221} = \Gamma_{122} = \Gamma_{212} = 0 \ , \ \Gamma_{222} = \frac{1}{2}\phi'(\beta) \ .$$

The components of the symmetric 3-tensor T are:

$$T_{111} = \frac{2\beta}{\mu^3} \ , \ T_{112} = T_{121} = T_{211} = -\frac{1}{\mu^2} \ ,$$

$$T_{221} = T_{122} = T_{212} = 0 \ , \ T_{222} = \phi'(\beta) \ .$$

The components of the α-connections are given by

$$\Gamma^{\alpha}_{111} = \frac{(1+\alpha)\beta}{\mu^3} \ , \ \Gamma^{\alpha}_{112} = \frac{\alpha-1}{2\mu^2} \ , \ \Gamma^{\alpha}_{121} = \Gamma^{\alpha}_{211} = \frac{1+\alpha}{2\mu^2} \ ,$$

$$\Gamma^{\alpha}_{122} = \Gamma^{\alpha}_{212} = \Gamma^{\alpha}_{221} = 0 \ , \ \Gamma^{\alpha}_{222} = \frac{1-\alpha}{2}\phi'(\beta) \ .$$

The sectional curvature of the Gamma manifold w.r.t. the α-connection is given by

$$k_{\alpha} = \frac{1-\alpha^2}{4} \frac{\phi(\beta) + \beta\phi'(\beta)}{4\mu^2\beta\phi(\beta)} \ .$$

We can clearly observe that the triple $(g, \nabla^{(1)}, \nabla^{(-1)})$ is a dually flat structure on the Gamma manifold.

5 Divergence functions

Consider a set $\mathcal{F}$ of objects such as 2D/3D images or probability distributions, etc. To measure the difference from one object to another in $\mathcal{F}$, one defines a function on the product $\mathcal{F} \times \mathcal{F}$, D, called divergence function, with the properties:

$$D(p,q) \geq 0 \; \forall \; p,q \in \mathcal{F} \text{ and } D(p,q) = 0 \Leftrightarrow p = q \; .$$

The number $D(p,q)$ measures the "gap" of p from q and is not in general reversible.

Examples 5.1

1. Let (M,d) a metric space. Then $D = \frac{1}{2}d^2$ is a divergence function on M. This divergence is reversible; i.e. $D(p,q) = D(q,p)$

2. Let $\Omega \subset \mathbb{R}^n$ be an open set subset and ψ a C^∞-function on Ω with $\frac{\partial^2 \psi}{\partial x^i \partial x^j}(x) > 0$, $\forall \; x \in \Omega$.
Then it holds

$$\psi(z) - \psi(x) - < z - x \; , \; \operatorname{grad} \psi(x) \geq 0 \; .$$

The function $D : \Omega \times \Omega \longrightarrow \mathbb{R}_+$ defined by

$$D(x,z) := \psi(z) - \psi(x) - \langle z - x \; , \; \operatorname{grad} \psi(x) \rangle$$

is a divergence function on Ω.

3. **f-divergences on probability distributions**

Let $\mathcal{P} = \mathcal{P}(X)$ be the space of probability distributions on a measure space $\mathcal{X}$.

$$\mathcal{P}(X) = \{ p : \mathcal{X} \longrightarrow \mathbb{R}_+ \; / \; \int_{\mathcal{X}} p(r) \; dr = 1 \} \; .$$

The space $\mathcal{P}$ is convex since for any $\lambda \in [0 \; , \; 1]$ and $p, q \in \mathcal{P}$, $\lambda p + (1 - \lambda) q \in \mathcal{P}$.

Let $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$ be a convex function with $f(1) = 0$ and $f''(1) = 1$. The function $D_f : \mathcal{P} \times \mathcal{P} \longrightarrow \mathbb{R}$ defined by

$$D_f(p,q) = \int_{\mathcal{X}} p(r) f(\frac{q(r)}{p(r))} \; dr$$

is a divergence function on $\mathcal{P}(X)$ called f-divergence.

A special family of f-divergences on $\mathcal{P}$ if given by the functions

$$f_\varrho(t) := \begin{cases} \frac{4}{1-\varrho^2}(\frac{1+t}{2} - t^{\frac{(1+\varrho)}{2}}) & \text{if} \varrho^2 \neq 1 \\ t \ln t & \text{if} \varrho = 1 \\ \ln(\frac{1}{t}) \text{if} \varrho = -1 & \end{cases}$$

For $\varrho = 0$ we get $f_0(t) = 4(\frac{1+t}{2} - \sqrt{t})$ and the corresponding divergence function $D_0 := D_{f_0}$ is called the Hellinger divergence. We have

$$D_0(p,q) = 4(1 - \int_{\mathbb{R}} \sqrt{p(r)q'r)}dr = 2\int_{\mathbb{R}} (\sqrt{p(r)} - \sqrt{q(r)})^2 dr$$

.
We can see that $d_0 p, q) := \sqrt{2D_0(p,q)}$ is a distance function on $\mathcal{P}(\mathbb{R})$ called Hellinger distance.

For $\varrho = -1$, $f_{-1}(t) = \ln(\frac{1}{t})$ and the corresponding divergence function $D_{-1} := D_{f_{-1}}$ is called Kullback-Leibler divergence. We have

$$D_{-1}(p,q) = \int_{\mathbb{R}} p(r) \ln \frac{p(r)}{q(r)} dr \ .$$

4. **α-divergences**

The α-divergence (also called $\alpha - \log$ divergence) on a statistical manifold $(M, g, \nabla^{(\alpha)})$ is locally defined by:

$$D^\alpha(p,q) \equiv D^\alpha(x(p), x(q)) := \frac{1}{2} g_{ij}(p) \Delta z^i \Delta z^j + \Gamma_{ijk} \Delta z^i \Delta z^j \Delta z^k \ ,$$

with $\Delta z^l = (x^l(p) - x^l(q))$, where $(x^l(p))_l$ and $(x^l(q))_l$ are the affine components of the point p and q.

The α-divergence recovers the original Riemannian metric g in the sense

$$\partial_i \partial_j D^\alpha(x(p), x(q))_{!p=q} = g_{ij}(p) \ .$$

Also from the relation between the dual connections $\nabla^{(\alpha)}$ and $\nabla^{(-\alpha)}$, we have:

$$D^\alpha(p,q) = D^{-\alpha}(q,p) \ .$$

For $\alpha = -\frac{1}{3}$, D^α is called the standard divergence.

6 Canonical divergence on a statistical manifold

Let (M, g, ∇, ∇^*) be a torsion-free and dually flat statistical manifold, $(u^1, \cdots, u^n)$ and $(v_1, \cdots, v_n)$ respectively the ∇-affine and the ∇^*-affine coordinates system, ψ and ϕ the associated potential functions.

The function $D : M \times M \longrightarrow \mathbb{R}$ defined by

$$D(p,q) := \psi(p) + \phi(q) - u^i(p) v_i(q)$$

is a divergence function on M called canonical divergence.

The function $D^* : M \times M \longrightarrow \mathbb{R}$ defined by

$$D^*(p, q) := \phi(p) + \psi(q) - u_i(q)v^i(p)$$

is also a divergence function on M called canonical dual divergence. It holds:

$$D^*(p, q) = D(q, p) \ .$$

We have

$$\partial_i D(p, q) = v_i(p) - v_i(q) \text{ and } \partial^j D(p, q) = u^j(q) - u^j(p) \ .$$

It follows that $\partial_i D(p, q)$ and $\partial^j D(p, q)$ vanish for all i at $p = q$.
Also

$$\partial_i \partial_j D(p, q)_{!p=q} = g_{ij}(p) \ \text{ and } \ \partial^i \partial^j D(p, q)_{!p=q} = g^{ij}(q) \ .$$

The canonical divergence is characterized by the relation

$$D(p, q) + D(q, r) - D(p, r) = (u^i(p) - u^i(q))(v_i(r) - v_i(q)) \ ,$$

which can be seen as a generalization of the cosine formula in Hilbert spaces,

$$\frac{1}{2}\|p - q\|^2 + \frac{1}{2}\|q - r\|^2 - \frac{1}{2}\|p - r\|^2 = < p - q \ , \ r - q > \quad (**).$$

From the precedent characterization of the canonical divergence, it follows that:

The ∇-geodesic from q to p is given by $tu^i(p) + (1 - t)u^i(p)$, $t \in [0\ ,\ 1]$, since $(u^i)_i$ are ∇-affine coordinates, and likewise, the ∇^*-geodesic from q to r is given by $tv_j(r) + (1 - t)v_j(q)$, $t \in [0\ ,\ 1]$.

If those two geodesics are orthogonal at q, then we have the Pythagoras relation:

$$D(p, r) = D(p, q) + D(q, r) \ .$$

In particular, such a point q where these two geodesics meet orthogonally is the point closest to p on that ∇^*-geodesic. We may therefore consider such a point q as the orthogonal projection of p on the that latter geodesic.

By the same argument we can characterize the orthogonal projection of p onto an autoparallel submanifold S of M by such a relation. We have

Proposition 6.1
Let S be a differentiable submanifold of M. Then $q \in S$ is a stationary point of the function $D_p = D(p, .) : r \in S \longmapsto D(p, r)$, if and and only if the ∇-geodesic from p to q meets S orthogonally.

Minima for the orthogonal projection onto a submanifold can then be characterize by the following:

Let S be a ∇^*-autoparallel submanifold of M and $p \in M$. Then a point $q \in S$ satisfies

$$q = \operatorname{argmin}_{r \in S} D(p, r)$$

if and only if the ∇-geodesic from p to q is orthogonal to S at q.

7 Warped products of dualistic structures

Several metrics on statistical models are represented as warped products like the Fisher metric on the multivariate Gaussian distributions space (the so-called Takano Gaussian space), the Bogoliubov-Kubo-Mori metrics, the denormalizations of the Fisher metric.

Denormalization: Let $\mathcal{S} = \{p(x, \eta) \ / \ x \in \mathcal{X} \ ; \ \eta \in \mathbb{R}^n\}$ be the set of probability distributions on a set $\mathcal{X}$, parameterized by $\eta \in \mathbb{R}^n$. It is an n-dimensional statistical manifold with the Fisher metric g. The denormalization of $\mathcal{S}$ is the space $\tilde{\mathcal{S}} = \{tp(x, \zeta) \ / \ t \in \mathbb{R}_{>0} \ , \ p(x, \zeta) \in \mathcal{S}\}$ of positive measures on $\mathcal{X}$, parameterized by $\eta \in \mathbb{R}^n$.

The space $\tilde{\mathcal{S}}$ could be modeled as the warped product $\mathbb{R}_{>0} \times \mathcal{S}$ with the metric $\tilde{g} = \frac{1}{t}\, dt^2 + t\, g$.

Takano Gaussian space:

$$\mathcal{M} = \{p(x, \zeta) \ / \ x \in \mathbb{R}^n \ ; \ \zeta \in \mathbb{R}_{>0} \times \mathbb{R}^n\}$$

with $p(x, \zeta) = \frac{1}{\sqrt{2\pi}\sigma} \Pi_{i=1}^n \exp\{-\frac{(x_i - m_i)^2}{2\sigma^2}\}$, where $\zeta = (\sigma, m_1, \cdots, m_n) \in \mathbb{R}_{>0} \times \mathbb{R}^n$. The Takano Gaussian space is an $(n+1)$-dimensional statistical manifold that can be seen as the product $\mathbb{R}_{>0} \times \mathbb{R}^n$ with the metric $\tilde{g} = \frac{2n}{\sigma^2} dt^2 + \frac{1}{\sigma^2} \delta_{ij}$.

Let $(M, g))$ and (N, h) be Riemannian manifolds of dimensions m and n respectively, $\pi : M \times N \to M$ and $\sigma : M \times N \to N$ be the canonical projections, and $f \in C^\infty(M)$ be a positive smooth function. The *warped product* of (M, g) and (N, h) with warping function f is the product manifold $M \times N$ endowed with metric tensor

$$G_f = \pi^* g + (f \circ \pi)^2 \sigma^* h \ ,$$

denoted by $G_f = g \oplus f^2 h$.

The tangent space $T_{(p,q)(M \times N)}$ at a point $(p, q) \in M \times N$, is isomorph to the direct sum $T_p M \oplus T_q N$.

Let $\mathcal{L}_H(M)$ (respectivement $\mathcal{L}_V(N)$) be the set of all vector fields on $M \times N$ which are horizontal lifts (respectivement vertical lifts) of vector fields on M (respectivement on N). We have

$$T(M \times N) = \mathcal{L}_H(M) \oplus \mathcal{L}_V(N)$$

and thus a vector field A on $M \times N$ can be written as

$$A = X + U, \quad \text{with} \quad X \in \mathcal{L}_H(M) \quad \text{and} \quad U \in \mathcal{L}_V(N).$$

Obviously

$$\pi_*(\mathcal{L}_H(M)) = TM \quad \text{and} \quad \sigma_*(\mathcal{L}_V(F)) = TN\ .$$

For any vector field $X \in \mathcal{L}_H(M)$, we denote $\pi_*(X)$ by $\bar{X}$ and for any vector field $U \in \mathcal{L}_V(N)$, we denote $\sigma_*(U)$ by $\tilde{U}$. Furthermore we denote the horizontal lift on $M \times N$ of a vector field $u \in TM$ by $(u)^{hor}$, and the vertical lift on $M \times N$ of a vector field $w \in TN$ by $(w)^{ver}$.

Let (G_f, D, D^*) be a dualistic structure on $M \times N$. For $X, Y \in \mathcal{L}_H(M)$ and $U, V \in \mathcal{L}_V(N)$, we put:

$$\pi_*(D_X Y) = {}^M\nabla_{\bar{X}}\bar{Y} \text{ and } \pi_*(D^*_X Y) = {}^M\nabla'_{\bar{X}}\bar{Y}$$

and

$$\sigma_*(D_U V) = {}^N\nabla_{\tilde{U}}\tilde{V} \text{ and } \sigma_*(D^*_U V) = {}^N\nabla'_{\tilde{U}}\tilde{V}\ .$$

${}^M\nabla$ and ${}^M\nabla'$ are affine connections on M, and ${}^N\nabla$ and ${}^N\nabla'$ are affine connections on N since D and D^* are affine connections on $M \times N$ and, $\pi : M \times N \longrightarrow M$ and $\sigma : M \times N \longrightarrow N$ the canonical projections.

Proposition 7.1
The triple $(g, {}^M\nabla, {}^M\nabla')$ is a dualistic structure on M and the triple $(h, {}^N\nabla, {}^N\nabla')$ is a dualistic structure on N; i.e.:

$${}^M\nabla' = {}^M\nabla^* \text{ w.r.t. } g \ \text{ and}$$

$${}^N\nabla' = {}^N\nabla^* \text{ w.r.t. } h\ .$$

Proof:
Let $\bar{X}, \bar{Y}, \bar{Z} \in TM$ and $X, Y, Z \in \mathcal{L}_H(M)$ their corresponding horizontal lifts respectively. Denoting the inner product w.r.t. G_f by $\langle\ ,\ \rangle$, we have:

$$\begin{aligned}
\bar{X} \cdot g(\bar{Y}, \bar{Z}) \circ \pi &= X \cdot \langle Y, Z \rangle \\
&= \langle D_X Y, Z \rangle + \langle Y, D^*_X Z \rangle \\
&= g(\pi_*(D_X Y), \pi_*(Z)) \circ \pi + g(\pi_*(Y), \pi_*(D^*_X Z)) \circ \pi \\
&= [g({}^M\nabla_{\bar{X}}\bar{Y}, \bar{Z}) + g(\bar{Y}, {}^M\nabla_{\bar{X}}\bar{Z})] \circ \pi\ .
\end{aligned}$$

Thus

$$\bar{X} \cdot g(\bar{Y}, \bar{Z}) = g({}^M\nabla_{\bar{X}}\bar{Y}, \bar{Z}) + g(\bar{Y}, {}^M\nabla_{\bar{X}}\bar{Z}) \ , \ \forall \ \bar{X}, \bar{Y}, \bar{Z} \in TM \ .$$

Hence ${}^M\nabla$ and ${}^M\nabla'$ are dual to each other w.r.t. g.
Also if $\tilde{U}, \tilde{V}, \tilde{W} \in TN$ and $U, V, W \in \mathcal{L}_V(N)$ their corresponding vertical lifts, we have:

$$\begin{aligned}
\tilde{U} \cdot h(\tilde{V}, \tilde{W}) \circ \sigma &= f^{-2} U \cdot \langle V, W\rangle \\
&= f^{-2}[\langle D_U V, W\rangle + \langle V, D^*_U W\rangle] \\
&= f^{-2}\{f^2 h(\sigma_*(D_U V), \sigma_*(W)) + f^2 h(\sigma_*(V), \sigma_*(D_U W))\} \circ \sigma \\
&= [h({}^N\nabla_{\tilde{U}}\tilde{V}, \tilde{W}) + h(\tilde{V}, {}^N\nabla'_{\tilde{U}}\tilde{W})] \circ \sigma \ ;
\end{aligned}$$

It follows then

$$\tilde{U} \cdot h(\tilde{V}, \tilde{W}) = h({}^N\nabla_{\tilde{U}}\tilde{V}, \tilde{W}) + h(\tilde{V}, {}^N\nabla'_{\tilde{U}}\tilde{W}) \ .$$

Hence ${}^N\nabla$ and ${}^N\nabla'$ are dual to each other w.r.t. h. □

In the following we construct a dualistic structure on the warped product $M \times_f N$ from those on its base manifold M and its fiber N.

Let (g, ∇, ∇^*) and $(h, \tilde{\nabla}, \tilde{\nabla}^*)$ be dualistic structures respectively on M and N. for $X, Y \in \mathcal{L}_H(M)$ and $U, V \in \mathcal{L}_V(N)$ we put:

$$\begin{aligned}
D_X Y &:= (\nabla_{\bar{X}}\bar{Y})^{hor} \\
D_X U = D_U X &:= \frac{X \cdot f}{f} U \\
D_U V &:= -\frac{\langle U, V\rangle}{f} \operatorname{grad}(f \circ \pi) + (\tilde{\nabla}_{\tilde{U}}\tilde{V})^{ver}
\end{aligned}$$

and

$$\begin{aligned}
D'_X Y &:= (\nabla^*_{\bar{X}}\bar{Y})^{hor} \\
D'_X U = D'_U X &:= \frac{X \cdot f}{f} U \\
D'_U V &:= -\frac{\langle U, V\rangle}{f} \operatorname{grad}(f \circ \pi) + (\tilde{\nabla}^*_{\tilde{U}}\tilde{V})^{ver}
\end{aligned}$$

Proposition 7.2
The triple (G_f, D, D') is a dualistic structure on $M \times N$; that is:

$$D' = D^* \text{ w.r.t. } G_f \ .$$

Proof
Let $X, Y, Z \in \mathcal{L}_H(M)$ and $U, V, W \in \mathcal{L}_V(N)$,

$$\begin{aligned}
X \cdot \langle X, Y\rangle &= \bar{X} \cdot g(\bar{Y}, \bar{Z}) \circ \pi \\
g(\nabla_{\bar{X}}\bar{Y}, \bar{Z}) \circ \pi + g(\bar{Y}, \nabla^*_{\bar{X}}\bar{Z}) \circ \pi &= \langle D_X Y, Z\rangle + \langle Y, D'_X Z\rangle \ ,
\end{aligned}$$

$$\begin{aligned}
U \cdot \langle V, W \rangle &= f^2 \tilde{U} \cdot h(\tilde{V}, \tilde{W}) \circ \sigma \\
&= f^2 [h(\tilde{\nabla}_{\tilde{U}} \tilde{V}, \tilde{W}) + h(\tilde{V}, \tilde{\nabla}_{\tilde{U}} \tilde{W})] \circ \sigma \\
&= \langle (\tilde{\nabla}_{\tilde{U}} \tilde{V})^{ver}, W \rangle + \langle V, (\tilde{\nabla}^*_{\tilde{U}} \tilde{W})^{ver} \rangle \\
&= \langle D_U V, W \rangle + \langle V, D'_U W \rangle \ ,
\end{aligned}$$

since $\operatorname{grad} f \in \mathcal{L}_H(M)$ and

$$\begin{aligned}
D_U V &= -\frac{\langle U, V \rangle}{f} \operatorname{grad} f + (\tilde{\nabla}_{\tilde{U}} \tilde{V})^{ver} \ , \\
D'_U V &= -\frac{\langle U, V \rangle}{f} \operatorname{grad} f + (\tilde{\nabla}^*_{\tilde{U}} \tilde{V})^{ver} \ .
\end{aligned}$$

Also

$$U \cdot \langle Y, Z \rangle = 0 = \langle D_U Y, Z \rangle + \langle Y, D'_U Z \rangle \ , \ \text{since} \ \langle D_U Y, Z \rangle = \langle Y, D'_U Z \rangle = 0 \ ,$$

$$X \cdot \langle V, Z \rangle = 0 = \langle D_X V, Z \rangle + \langle V, D'_X Z \rangle \ , \ \text{since} \ \langle D_X V, Z \rangle = \langle V, D'_X Z \rangle = 0 \ ,$$

$$X \cdot \langle Y, W \rangle = 0 = \langle D_X Y, W \rangle + \langle Y, D'_X W \rangle \ , \ \text{since} \ \langle D_X Y, W \rangle = \langle Y, D'_X W \rangle = 0 \ .$$

It follows from the relations above that:

$$A \cdot \langle B, C \rangle = 0 = \langle D_A B, WC \rangle + \langle B, D'_A C \rangle \ , \ \forall \ A, B, C \in T(M \times N) \ .$$

Hence D and D' are dual, that is $D' = D^*$, w.r.t. G_f □

We call the triple (G_f, D, D^*) the dualistic structure on $M \times N$ induced from (g, ∇, ∇^*) on M and $(h, \tilde{\nabla}, \tilde{\nabla}^*)$ on N.

Note that if the connections ∇ , ∇^*, $\tilde{\nabla}$, $\tilde{\nabla}^*$ are torsion-free, then the induced connections D and D^* are also torsion-free.
From now on all connections are assumed to be torsion-free.

Let ${}^gR$, ${}^hR$ and R be the Riemannian curvature operators w.r.t. ∇, $\tilde{\nabla}$ and D respectively. For $X, Y, Z \in \mathcal{L}_H(M)$ and $U, V, W \in \mathcal{L}_V(N)$, it holds:

$$R_{XY} Z = ({}^g R_{\bar{X}\bar{Y}} \bar{Z})^{hor} \ , \ R_{VY} Z = \frac{1}{f} H_f(Y, Z) V \ ,$$

$$R_{XY} U = 0 = R_{VW} Z \ , \ R_{XV} W = \frac{\langle V, W \rangle}{f} D_X(\operatorname{grad} f) \ ,$$

$$R_{VW} U = ({}^h R_{\tilde{V}, \tilde{W}} \tilde{U})^{ver} - \frac{\langle \operatorname{grad} f, \operatorname{grad} f \rangle}{f^2} (\langle V, U \rangle W - \langle W, U \rangle V) \ .$$

When the warping function f is constant it follows from the previous curvature relations that the warped product space $(M \times N, G_f, D, D^*)$ is dually flat if and only if (M, g, ∇, ∇^*) and $(N, h, \tilde{\nabla}, \tilde{\nabla}^*)$ are dually flat.

When the warping function f is non-constant, we have the following result:

Theorem 7.1:
Assume M to be connected. Then
$(M \times N, G_f, D, D^*)$ is dually flat if and only if, (M, g, ∇, ∇^*) is dually flat, $\operatorname{grad} f$ is a parallel vector field and (N, h) is a Riemannian manifolds of constant $\tilde{\nabla}$-sectionnal curvature $k = \| \operatorname{grad} f\|^2$.

Proof
If (g, ∇, ∇^*) is flat on M, $\operatorname{grad} f$ is a parallel vector field and (N, h) is a Riemannian manifolds of constant sectionnal curvature $k = \| \operatorname{grad} f\|^2$, it is immediate to get the result. Assume (G_f, D, D^*) to be dually flat on $M \times N$.
Then from the relation $R_{XY}Z = ({}^g R_{\bar{X}\bar{Y}}\bar{Z})^{hor}$ we have ${}^g R = 0$; i.e. ∇ is a flat connection.
Since the connections are torsion-free, its dual connection ∇ is also flat and therefore (G_f, ∇, ∇^*) is a flat dualistic structure on $M \times N$.
For $X \in \mathcal{L}_H(M)$ and $V, W \in \mathcal{L}_V(N)$, we have $R_{XV}W = \frac{\langle V,W\rangle}{f} D_X(\operatorname{grad} f) = 0$ and then

$$D_X(\operatorname{grad} f) = 0 \ , \ \forall \ X \in \mathcal{L}_H(M) \ .$$

Thus $\operatorname{grad} f$ is a parallel vector field and then of constant norm.
Moreover for $U, V, W \in \mathcal{L}_V(N)$, we have $R_{VW}U = ({}^h R_{\tilde{V},\tilde{W}}\tilde{U})^{ver} - \frac{\langle \operatorname{grad} f, \operatorname{grad} f\rangle}{f^2}(\langle V, U\rangle W - \langle W, U\rangle V) = 0$.
Hence

$${}^h R_{\tilde{V},\tilde{W}}\tilde{U} = \| \operatorname{grad} f\|^2 \{h(\tilde{V}, \tilde{U})\tilde{W} - h(\tilde{W}, \tilde{U})\tilde{V}\} \ , \ \forall \ U, V, W \in \mathcal{L}_V(N) \ .$$

Since $\| \operatorname{grad} f\|$ is constant, it follows that (N, h) has a constant $\tilde{\nabla}$-sectionnal curvature equal to $\| \operatorname{grad} f\|^2$. □.

7.1 Case of cones

A cone is a warped product of $\mathbb{R}_{>0}$ with a Riemannian manifold (N, h); i.e.: $\mathbb{R}_{>0} \times_f N$ with $f : \mathbb{R}_{>0} \longrightarrow \mathbb{R}_{>0}$.

Let us consider flat dualistic structures (G_f, D, D^*) on $\mathbb{R}_{>0} \times N$ and $(h, \tilde{\nabla}, \tilde{\nabla}^*)$ on N. If (g, ∇, ∇^*) a dually flat structure on $\mathbb{R}_{>0}$ and $f(t) = t$, $\forall \ t \in \mathbb{R}_{>0}$, then

$$\nabla_{\partial_t}\partial_t = \frac{1}{t}\partial_t \ \text{ and } \ \nabla^*_{\partial_t}\partial_t = -\frac{1}{t}\partial_t \ \text{ with } \ \partial_t = \frac{\partial}{\partial t} \ .$$

7.2 Case of twisted products

The twisted product of the pseudo-Riemannian manifolds (M, g) and (N, h) is the cartesian product $M \times N$ endowed with the metric

$$G_f = \pi^* g + f^2 \sigma^* h \ ,$$

where the warping function $f : M \times N \longrightarrow \mathbb{R}_{>0}$ is defined on the $M \times N$ and not only on M as for the warped product.

The projections of a connection of $M \times_f N$ on M and N are defined as in the case of warped product, but the connection D on $M \times_f N$ induced by a connection ∇ on M and a connection $\tilde{\nabla}$ on N is defined for $X, Y \in \mathcal{L}(B)$ and $U, V \in \mathcal{L}(F)$ by:

$$\begin{aligned} D_X Y &= \nabla_{\bar{X}} \bar{Y}; \\ D_X U = D_U X &= (X \cdot \phi) U; \\ D_U V &= (\tilde{\nabla}_{\tilde{U}} \tilde{V})^{ver} + (U \cdot \phi) V + (V \cdot \phi) U - h(\tilde{U}, \tilde{V}) \operatorname{grad} \phi \end{aligned}$$

with $\phi = \log f$.

Torsion-free dualistic structures (∇, ∇^*) on (M, g) and $(\tilde{\nabla}, \tilde{\nabla}^*)$ on (N, h) induce torsion-free dualistic structures (D, D^*) on the twisted product $(M \times N, G_f)$ and conversely the projections of a torsion-free dualistic structure (D, D^*) on $(M \times N, G_f)$ are also torsion-free dualistic structures on (M, g) and (N, h).

Concerning the flatness of twisted dualistic structures, we have the following result:

Proposition 7.3
Assume that *dim* $N > 1$ and $\mathrm{Ric}(X, U) = 0\,; \forall\, X \in \mathcal{L}_H(M)$ and $\forall\, U \in \mathcal{L}_V(N)$. Then

$(M \times N, G_f, D, D^*)$ is dually flat, if and only if (M, g, ∇, ∇^*) is dually flat and (N, h) is a conformal to a Riemannian manifold of constant sectionnal curvature.

References

[1] S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics, 28, Springer-Verlag, 1985.

[2] S. Amari and H. Nagaoka, *Methods of Information geometry*, AMS, Oxford University Press, vol. 191, 2000.

[3] N. Ay, J Jost, H. V. Lê, L. Schwachhöfer, *Information geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 64, Springer Verlag, 2017. ISBN 978-3-319-56477-7 ; DOI 10.1007/978-3-319-56478-4

[4] N. Ay and W. Tuschmann, *Dually flat manifolds and global information geometry*, Open Syst. and Information Dyn. 9 (2002), 195-200.

[5] A. S. Diallo, L. Todjihounde, *Dualistic structures on twisted product manifolds*, Global Journal of Advanced Research on Classical and Modern Geometries, Vol.4, (2015), Issue 1, pp.35-43.

[6] Y. Fujitani, *Information geometry of warped product spaces*, arXiv:2202.02762.v2, 2022.

[7] B. O'Neill, *Semi-Riemannian geometry*, Academic Press, New-York, 1983. https://fr.overleaf.com/project/6364fbbb1027f21c55a61b77

[8] S. L. Lauritzen, *Statistical manifolds*, Institute of Mathematical Statistics Lecture Notes - Monograph Series, 1987: 163-216 (1987). https://doi.org/10.1214/lnms/1215467061.

[9] S. Said, L. Bombrun, Y. Berthoumieu, *Warped Riemannian metrics for location-scale models*, Geometric structures of information, 251-296 (2019).

[10] A. Takatsu, Wasserstein geometry of Gaussian measures, Osaka J. Math. 48 (2011), no. 4, 1005-1026.

[11] L. Todjihounde, *Dualistic structures on warped product manifolds*, Diff. Geom.-Dyn. Syst. 8, (2006), 278-284.