Geometry of tangent and cotangent bundles on statistical manifolds

Tat Dat Tô, Stéphane Puechmorel

Geometry of tangent and cotangent bundles on statistical manifolds

Tat Dat Tô and Stéphane Puechmorel

Ecole Nationale de l'Aviation Civile, Unversité Fédérale de Toulouse
7, Avenue Edouard Belin, FR-31055 Toulouse, France
{tat-dat.to, stephane.puechmorel}@enac.fr

Abstract. This note surveys some results on the geometric structure on the tangent bundle and cotangent bundle of statistical manifolds. We also study the completeness of the tangent bundle with respect to the Sasaki metrics induced from the statistical structures.

Keywords: Statistical manifolds · Hessian manifolds · almost Kähler structure.

1 Statistical manifolds

We recall here the definition of statistical manifolds following [8, 2, 1, 3]. Let (M, g) be a Riemannian manifold. Denote $\Gamma(TM)$ the space of vector fields on M. A *connection* ∇ on M is a bilinear map

$$\nabla : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM) \tag{1}$$

$$(X, Y) \mapsto \nabla_X Y \tag{2}$$

satisfying for all $f \in C^\infty(M), X, Y, Z \in \Gamma(TM)$

1. $\nabla_{fX} Y = f\nabla_X Y$, i.e, ∇ is $C^\infty(M, \mathbf{R})$-linear in the first variable;
2. $\nabla_X(fY) = X(f)Y + f\nabla_X Y$, i.e, ∇ satisfies *Leibniz rule* in the second variable.

A connection ∇^* is called *dual connection* of ∇ with respect to g if

$$\nabla_Z g(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla^*_Z Y)$$

for all $X, Y, Z \in \Gamma(TM)$. It is straightforward that ∇^* is defined uniquely by the last identity.

Definition 1. *A triple (M, g, ∇) is called a statistical manifold if ∇ and ∇^* are torsion free. We also call (g, ∇, ∇^*) a dualistic structure on M.*

Assume that (g, ∇, ∇^*) is a dualistic structure on M. Then the tensor T defined by

$$T(X, Y, Z) = g(\nabla^*_X Y - \nabla_X Y, Z), \forall X, Y, Z \in \Gamma(TM)$$

is symmetric on all three entries. The tensor T has been called the skewness tensor by Lauritzen [8]. Conversely, it follows from Lauritzen [8] that:

Theorem 1. *A metric g and a symmetric 3-tensor T yields a dualistic structure with torsion-free connections.*

Denote $\hat{\nabla}$ the Levi-Civita connection with respect to g, then we have

$$\hat{\nabla} = \frac{1}{2}(\nabla + \nabla^*) \quad \text{and} \quad T = 2(\hat{\nabla} - \nabla).$$

From now we have another way to define a statistical manifold by a triple (M, g, T), where g is a Riemannian metric on M and T is a symmetric 3-tensor.

Definition 2. *A smooth map ϕ from a statistical manifold (M_1, g_1, T_1) to a statistical manifold (M_2, g_2, T_2) is called a statistical isomorphism if ϕ is an diffeomorphism of M_1 into M_2 such that $g_1 = \phi^*(g_2)$ and $T_1 = \phi^*(T_2)$.*

2 Tangent and Cotangent bundles of statistical manifolds

2.1 Almost Kähler structure on tangent bundles

We recall here the work of Dombrowski [5] on the geometry of the tangent bundle. As a corollary, the tangent bundles of a statisticcal manifolds admit a canonical almost Kähler structure.

Let (M, g) be a Riemannian manifold and ∇ be a connection on M. Denote by $\pi : TM \to M$ the canonical projection, then $\pi_* := d\pi : TTM \to TM$ is a C^∞ map. We denote by $V_\xi TM$ the vector space $\ker(\pi_{*|_{T_\xi TM}})$. Then the connection ∇ induces a direct sum decomposition on TTM at any $\xi \in TM$:

$$T_\xi TM = V_\xi TM \oplus H_\xi TM \tag{3}$$

where $V_\xi TM$ (reps. $H_\xi TM$) is called the vertical (reps. horizontal) space of $T_\xi TM$. Indeed, for any a normal neighborhood U (with respect to ∇) of $p \in M$, there is a canonical map

$$\tau : \pi^{-1}(U) \to T_p M$$

defined as follows: for $\xi \in \pi^{-1}(U)$, $\tau(\xi)$ is the parallel translation with respect to ∇ of ξ along the unique geodesic arc in U joining $q = \pi(\xi)$ and p. We define the *connection map* $K : TTM \to TM$ as follows: for any $\xi \in TM$, $A \in T_\xi TM$, and $c : t \to c(t)$ a path in TM with $\dot{c}(0) = A$, then

$$KA = \lim_{t \to 0} \frac{\tau(c(t)) - \xi}{t}.$$

Then we define $H_\xi TM := \ker(K_{|_{T_\xi TM}})$.

Let $u^1, \ldots, u^n \in C^\infty(M)$ be a coordinate system on M. Define then a coordianate system $v^1, \ldots, v^{2n} \in C^\infty(TM)$ on TM as follows:

$$v^i := u^i \circ \pi \quad \forall i = 1, \ldots, n \tag{4}$$

$$v^{n+i}(X) := du^i(X) \quad \forall i = 1, \ldots, n, \text{ and } \forall X \in TM. \tag{5}$$

We also denote

$$X_i = \frac{\partial}{\partial u^i} \quad and \quad A_j = \frac{\partial}{\partial v^j}, \tag{6}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, 2n$. Then for any $\xi = \sum_{i=1}^n x^i X_i \in TM$ with $x^i \in C^\infty(M)$, and $A = \sum_{j=1}^{2n} a^j A_j \in \Gamma(TM)$ with $a^j \in C^\infty(TM)$, we have

$$(\pi_* A)_\xi = \sum_{i=1}^n a^i(\xi) X_i$$

and

$$(KA)_\xi = \sum_{i,j,k=1}^n (a^{n+i}(\xi) + \Gamma^i{}_{jk} a^j(\xi) x^i) X_i, \tag{7}$$

where $\Gamma^i{}_{jk} = (\nabla_{X_j} X_k) u^i$.

We now define an almost complex structure J on TM as follows. For $\xi \in TM$, the exists for $A \in T_\xi TM$ a unique element in $T_\xi TM$ denoted by JA such that $\pi_*(JA) = -KA$ and $K(JA) = \pi_* A$. The map $J : TTM \to TM$ is thus an almost complex structure for TM characterized by

$$\pi_* \circ J = -K, \quad K \circ J = \pi_*. \tag{8}$$

It follows from [5] that for any $X \in \Gamma(TM)$, there exist unique $X^h \in \Gamma(TTM)$, called *horizontal lift* and unique $X^v \in \Gamma(TTM)$, called *vertical lift* satisfying

$$\pi_*(X^h_\xi) = X_{\pi(\xi)}, \pi_*(X^v_\xi) = O_{\pi(\xi)}, KX^h_\xi = O_{\pi(\xi)}, KX^v_\xi = X_{\pi(\xi)}.$$

By the definition we have $JX^h = X^v$ and $JX^v = -X^h$ for ant $X \in \Gamma(TM)$.

We also have a natural Riemannian metric $\tilde{g}$ on TM, namely *Sasaki metric* (cf. [10, 5]) induced from (g, ∇):

$$\tilde{g}(A, B) := g(\pi_* A, \pi_* B) + g(KA, KB), \quad \forall A, B \in T_\xi TM. \tag{9}$$

Observe J is g-compatible, i.e, $\tilde{g}(JA, JB) = \tilde{g}(A, B)$, therefore we can define a hermitian metric $\tilde{h}$ and its Kähler form $\tilde{\omega}$ on TTM by

$$\tilde{h}(A, B) = \tilde{g}(A, B) + i\tilde{g}(A, JB), \quad \forall A, B \in T_\xi TM, \tag{10}$$

and

$$\tilde{\omega}(A, B) = \tilde{h}(A, JB), \quad \forall A, B \in T_\xi TM. \tag{11}$$

Then it follows from Dombrowski [5] and Satoh [11] that:

Theorem 2. *Let (M, g) be a Riemannian manifold with a torsion-free connection ∇, then $(TM, \tilde{h}, J)$ is an almost-Hermitian manifold. It is almost-Kählerian if and only if ∇^* is torsion-free (so that (M, g, ∇) is statistical). Furthermore, $(TM, \tilde{h}, J)$ is a Kähler manifold if and only if ∇ is flat, i.e the Riemannian curvature of ∇ vanishes.*

This theorem gives a bridge between information geometry and complex geometry. As consequence we have the following characterization of statistical isomorphism.

Theorem 3. *A smooth map $\phi : (M_1, g_1, T_1) \to (M_2, g_2, T_2)$ on statistical manifolds is a statistical isomorphism if and only if $\varphi := d\phi : (TM_1, \tilde{h}_1, J_1) \to (TM_2, \tilde{h}_2, J_2)$ is an isomorphism of almost-Kähler manifolds, i.e, φ is a diffeomorphism satisfying $\varphi^* \tilde{h}_2 = \tilde{h}_1$ and $\varphi^* J_2 = J_1$.*

Proof. The complex structure on TM_1 is defined by g_1 and $\nabla^1 = \hat{\nabla}^1 - \frac{1}{2} T_1$, where $\hat{\nabla}^1$ is the Levi-Civita connection of g_1. The complex structure on TM_2 is defined in the same way.

Suppose that $\phi : (M_1, g_1, T_1) \to (M_2, g_2, T_2)$ is a statistical isomorphism. Then it follows from the definition that $g_1 = \phi^* g_2$ and $T_1 = \phi^* T_2$, hence we get $\varphi^* \tilde{h}_2 = \tilde{h}_1$ and $\varphi^* J_2 = J_1$.

Conversely, if $\varphi := d\phi : (TM_1, \tilde{h}_1, J_1) \to (TM_2, \tilde{h}_2, J_2)$ is an isomorphism of almost-Kähler manifolds, then ϕ is diffeomorphic. Since $\varphi^* \tilde{h}_2 = \tilde{h}_1$, the formulas (10) and (9) imply that $\phi^*(g_2) = g_1$. Finally, the identities (8) and $\varphi^* J_2 = J_1$ imply $\phi^* K_2 = K_1$, where K_1, K_2 are the connections maps of $(M_1, g_1, T_1), (M_2, g_2, T_2)$. Since the connection ∇^1, ∇^2 can be defined by K_1, K_2 (see (7)), so do T_1 and T_2. This show that $\phi^* T_2 = T_1$, hence ϕ is a statistical isomorphism.

2.2 Cotangent bundle

Let M be a manifold, then its cotangent bundle T^*M has a natural symplectic structure defined as follows. Denote by $\bar{\pi} : T^*M \to M$ the canonical projection which assigns to each form $p \in T_q^*M$ its base point $q \in M$. Define the *Liouville 1-form* θ on T^*M by

$$\langle \theta, X \rangle = \langle p, \bar{\pi}_* X \rangle, \forall X \in T_p T^*M. \tag{12}$$

Then $\omega = d\theta$ is a symplectic form on M. Take a local coordinate system $(p_1, \ldots, p_n, q_1, \ldots, q_n)$ on T^*M, where $(q_1, \ldots, q_n)$ is a local coordinate on M and $(p_1, \ldots, p_n)$ are the coefficients of forms. Then $\theta = \sum p_j dq_j$ and $\omega = -d\theta = dq_j \wedge dp_j$.

Now let g be a metric on M then, we have the following isomorphism

$$\phi_g : TM \to T^*M \tag{13}$$

$$\xi \mapsto g(\cdot, \xi) \tag{14}$$

The following result is straightforward from Delanoë [4, Théorème 2] and Satoh [11].

Theorem 4. *Let (M, g) be a Riemannian manifold with a torsion-free connection ∇. The following are equivalent:*

1. *(M, g, ∇) is a statistical manifold;*
2. *$(\phi_g)_*(HM)$ is a Lagrangian subspace with respect to the symplectic form ω on T^*M, i.e, ω vanishes on $(\phi_g)_*(HM)$;*
3. *$\phi_g : (TM, \tilde{\omega}) \to (T^*M, \omega)$ satisfies $\phi_g^* \omega = \tilde{\omega}$, where $\tilde{\omega}$ defined in (11).*

2.3 Käher and dual flat correspondence

Definition 3. *A statistical manifold* (M, g, ∇) *is said to be* ∇*-flat if* ∇ *is a flat connection.*

Since ∇ being flat implies that ∇^* is flat as well, we call a ∇-flat statistical manifold is *dual flat*. It follows from [2] that any dual flat statistical manifold (M, g, ∇) carries a Hessian structure, i.e g can be locally expressed by $g = \nabla d\varphi$, that is,

$$g_{ij} = \frac{\partial^2 \varphi}{\partial x^i \partial x^j}$$

where $\{x^1, \ldots, x^n\}$ is an affine coordinate system with respect to ∇. By Theorem 2, $(TM, \tilde{g}, J)$ is a Kähler manifold, where $\tilde{g}$ and J are defined in Section 2.1.

Denote $\hat{\nabla}$ the Levi-Civita of (M, g), $T = \frac{1}{2}(\hat{\nabla} - \nabla)$. Then the (1,3) tensor $Q = \nabla T$ is called the Hessian curvature tensor for (g, ∇).

Definition 4. *We define first Koszul form* α *and the the second Koszul form* β *for* (g, ∇) *(cf. [7]) by*

$$\nabla_X vol_g = \alpha(X) vol_g \quad \text{and} \quad \beta = -\nabla\alpha.$$

We recall here some properties of the Hessian curvature (cf. [12])

Proposition 1. *Let* $\hat{R}$ *be the Riemannian curvature of* g*,* $\tilde{R}$ *be the Riemannian curvature on the Kähler manifold* $(TM, \tilde{g}, J)$*. Then*

$$\hat{R}_{ijkl} = \frac{1}{2}(Q_{ijkl} - Q_{jikl}), \quad \tilde{R} = Q \circ \pi, \quad \tilde{R}_{j\bar{k}} = \beta_{jk} \circ \pi,$$

where $\tilde{R}_{j\bar{k}}$ *is the Ricci tensor on* $(TM, \tilde{g}, J)$*.*

These properties above infer that the second Koszul form β plays a similar role as the Ricci tensor in Kähler geometry.

3 Completeness to tangent bundles of statistical manifolds

In this section we study the completeness to the tangent bundle $(TM, \tilde{g})$ of a statistical manifold (M, g, ∇). We first recall the Riemannian submersion.

Definition 5. *Let* $(\tilde{M}, g)$ *and* (M, h) *be two Riemannian manifolds and* $\pi : \tilde{M} \to M$ *be a submersion. Then* π *is call* Riemannian submersion *if the isomorphism* $\pi_* : H\tilde{M} \to TM$ *is an isometry, where* $H\tilde{M}$ *is the horizontal distribution.*

We now have the following lemma (see for example [9]).

Lemma 1. *Let* $\pi : \tilde{M} \to M$ *be a Riemannian submersion. Then we have that*

(i) for any gedesic γ *of* M*,* $\tilde{p} = \pi^{-1}(\gamma(0))$*, there exists a unique locally defined horizontal lift* $\tilde{\gamma}$ *of* γ *with* $\tilde{\gamma}(0) = \tilde{p}$*, and* $\tilde{\gamma}$ *is a geodesics of* $\tilde{M}$*.*

(ii) *for any geodesic* $\tilde{\gamma}$ *of* $\tilde{M}$. *If* $\tilde{\gamma}'(0)$ *is a horizontal vector, then* $\tilde{\gamma}'(t)$ *is horizontal for every* t *in the domain of* $\tilde{\gamma}$ *and* $\pi \circ \tilde{\gamma}$ *is a geodesic of* M *of the same length as* $\tilde{\gamma}$.

(iii) *If* $\tilde{M}$ *is complete, so is* M.

We now prove that a necessary condition so that $(TM, \tilde{g})$ is complete is that (M, g) is complete.

Proposition 2. *If* $(TM, \tilde{g})$ *is complete, so is* (M, g).

Proof. We only need to prove that the canonical map $\pi : (TM, \tilde{g}) \to (Mg)$ is a Riemannian submersion. Indeed, it follow from the definition of Sasaki metric $\tilde{g}$ that

$$\tilde{g}(A, B) := g(\pi_* A, \pi_* B) + g(KA, KB), \quad \forall A, B \in T_\xi TM.$$

Therefore, for any $A, B \in H_\xi TM$ we have

$$\tilde{g}(A, B) := g(\pi_* A, \pi_* B)$$

this implies that π is a Riemannian submersion. Lemma 1 now implies that (M, g) is complete.

Theorem 5. *Suppose* (M, g, ∇) *is a complete statistical manifold, then* $(TM, \tilde{g})$ *is also complete.*

Proof. For any $p \in TM$ and any $\xi \in T_pTM$, the Cauchy-Lipschitz theorem implies the existence of the geodesic $\tilde{\gamma}$ on an open interval $[0, t_0)$, starting from p with the tangent vector ξ on $(TM, \tilde{g})$. We can assume $\tilde{\gamma} : [0, t_0) \to TM$ is a unit-speed geodesic. In order to prove that $(TM, \tilde{g})$ is complete, we need to prove that the geodesic is defined on the interval $[0, t_0]$ as well.

We now prove that, $\tilde{\gamma}$ can be extended beyond t_0. Indeed, take $\gamma = \pi_\circ \tilde{\gamma}$, then γ is Lipschitz since $\tilde{\gamma}$ has bounded speed. In addition, (M, g) is complete, hence there exists $p = \lim_{t \to t_0^-} \gamma(t)$. Take a neighborhood U of p, then $\pi^{-1}(U) = U \times \mathbf{R}^n$ and $\tilde{\gamma}(t) = (\gamma(t), c(t)) \in \pi^{-1}(U)$ for $t \in (t_0 - \epsilon, t_0)$ for some $\epsilon > 0$, hence $c((t_0 - \epsilon, t_0)) \subset \mathbf{R}^n$. Since the speed of $c(t)$ is also bounded and the restriction of $\tilde{g}$ on fibers is complete, there exists $\lim_{t \to t_0^-} c(t)$, therefore we can extend $\tilde{\gamma}$ until $t = t_0$.

Corollary 1. *Suppose that* (M, g, ∇) *is a compact statistical manifold, then* $(TM, \tilde{g})$ *is complete.*

4 Examples and applications

4.1 Euclidean space

Let $M = \mathbf{R}^n$, ∇ be from the standard derivative and

$$\psi = \frac{1}{2} \sum_{i=1}^{n} (x_i)^2, \quad \text{then} \quad g = \sum_{i=1}^{n} (dx^i)^2$$

is the standard Euclidean metric. Then $T\mathbf{R}^n$ is identified with $\mathbf{C}^n$. Then $(TM, J, \tilde{g})$ is a complete manifold.

4.2 Poincaré metric model

Let $M = \mathbf{R}^+ = \{x \in \mathbf{R} | x > 0\}$ and $\psi = \log(x^{-1}), \forall x \in \mathbf{R}^+$. We then have $g = \frac{1}{x^2} dx^2$. Then the tangent space $T\mathbf{R}^+ = \{(x, y) | x > 0\}$ has the induced metric

$$\tilde{g} = \frac{(dx)^2 + (dy)^2}{x^2}.$$

Therefore $(T\mathbf{R}^+, \tilde{g})$ is the Poincaré half-plane model and $\tilde{g}$ is the Poincaré metric. It is known that the Poincaré half-plane model is complete.

4.3 Regular convex cone

Let Ω is a regular convex cone (cf. [12]) and Ω^* be its dual cone. Denote by ψ the characteristic function on Ω

$$\psi(\theta) = \int_{\Omega^*} e^{-\langle x, \theta \rangle} dx. \tag{15}$$

For $x \in \Omega^*$, and $\theta \in \Omega$, define

$$p(x; \theta) = \frac{e^{-\langle x, \theta \rangle}}{\psi(\theta)} = e^{-\langle x, \theta \rangle - \log \psi(\theta)}. \tag{16}$$

Then $\{p(x; \theta) | \theta \in \Omega\}$ is an exponential family of probability distributions on Ω^* parametrized by Ω and its statistical structure $(\Omega, \nabla, g = \nabla d \log \psi)$ is a dual flat, where ∇ defined by ${\Gamma^i}_{jk} = g^{is} E_\theta \left[\partial_j \partial_k (\log p) \partial_s (\log p)\right]$. Then we have the following result.

Theorem 6. *Suppose Ω is homogeneous regular convex cone, then the Kähler metric $\tilde{h}$ (see (10)) is the Bergman metric on $T\Omega := \Omega + i\mathbf{R}^n$ and $(T\Omega, \tilde{g})$ is complete.*

Proof. Let ψ^* be characteristic function for the convex cone Ω^*:

$$\psi^*(x) = \int_{\Omega} e^{-\langle x, \theta \rangle} d\theta. \tag{17}$$

Then it follows from [6] that the Bergman kernel K_Ω of $T\Omega := \Omega + i\mathbf{R}^n$ is defined by

$$K_\Omega(z, w) = \pi^{-n} \int_{\Omega^*} e^{-\langle x, z + \bar{w} \rangle} \psi^*(x)^{-1} dx. \tag{18}$$

Denote $g = \nabla d \log \psi$ is the Fisher metric of (Ω, g, ∇). We now prove that $\tilde{h}$ is the Bergman metric $\frac{i}{2} \partial \bar{\partial} \log K_\Omega(z, z)$ on $T\Omega := \Omega + i\mathbf{R}^n$. Indeed, we have

$$K_\Omega(z, z) = \pi^{-n} \int_{\Omega^*} e^{-2\langle x, \theta \rangle} \psi^*(x)^{-1} dx =: h(\theta) \tag{19}$$

where $\theta = Re(z)$. Observe that for any $A \in Aut(\Omega)$, $h(A\theta) = (\det(A))^{-2} h(\theta)$ and $\psi(\theta) = (\det A)^{-1} \psi(\theta)$. For a fixed $\theta_0 \in \Omega$, the homogeneity of Ω implies that for any $\theta \in \Omega$, there exists $A \in Aut(\Omega)$ such that $\theta = A\theta_0$. Therefore if $h(\theta_0) = c\psi^2(\theta_0)$ for some c, then $h(\theta) = c\psi^2(\theta), \forall\theta \in \Omega$. Since $g = Dd\log\psi$ and $\tilde{h}_{j\bar{k}} = g_{jk} \circ \pi$, we have

$$\tilde{g}_{j\bar{k}}(z) = \partial_j\partial_k \log\psi(Re(z)) = \frac{1}{2}\partial_j\partial_{\bar{k}} \log K_{\Omega}(z, z).$$

Since (Ω, g) is a homogeneous manifold, it is complete. The conclusion is now followed from Theorem 5.

References

1. Amari, S.: Information Geometry and Its Applications. Applied Mathematical Sciences, vol. 194. Springer, Berlin (2016)
2. Amari, S., Nagaoka, H.: Methods of Information Geometry. Translations of Mathematical Monographs, vol. 191. Am. Math. Soc./Oxford University Press, Providence/London (2000)
3. Ay, N., Jost, J., Le, H.V., Schwachhöfer, L.: Information Geometry, Springer, (2017)
4. Delanoë, P.: Remarques sur les variétés localement hessiennes. Osaka J. Math. 26 (1989), no. 1, 65–69.
5. Dombrowski, P.: On the geometry of the tangent bundle. J. Reine Angew. Math. 210 (1962) 73–88.
6. Korányi, A.: The Bergman kernel function for tubes over convex cones. Pacific J. Math. 12 (1962), no. 4, 1355–1359.
7. Koszul, J-L.: Domaines bornés homogènes et orbites de groupes de transformations affines. Bulletin de la S. M. F., tome 89 (1961), p. 515-533.
8. Lauritzen, S.: Statistical manifolds. Differential geometry in statistical inference, 10:163– 216, 1987.
9. Petersen, P.: Riemannian Geometry. third edition, GTM 171, Springer (2016)
10. Sasaki. S.: On the differential geometry of tangent bundle of Riemannian manifolds. Tôhoku Math. J.,10 (1958), 338–354.
11. Satoh, H.: Almost Hermitian structures on tangent bundles. Proceedings of the Eleventh International Workshop on Differential Geometry, 105–118, Kyungpook Nat. Univ., Taegu, 2007.
12. Shima, H.: The Geometry of Hessian Structures. World Scientific Pub Co Inc (March 1, 2007)