Métriques naturelles sur les fibrés en géométrie de l'information

Florence Nicol, Stéphane Puechmorel

Métriques naturelles sur les fibrés en géométrie de l'information

Florence Nicol [1] & Stéphane Puechmorel [2]

[1] *Université de Toulouse, ENAC, 7 Avenue Edouard Belin, 31055 Toulouse, nicol@recherche.enac.fr*
[2] *Université de Toulouse, ENAC, 7 Avenue Edouard Belin, 31055 Toulouse, stephane.puechmorel@enac.fr*

Résumé. La géométrie de l'information permet une formulation élégante de propriétés statistiques. De nombreux travaux ont été dédiés à l'étude des relations entre représentation géométrique et estimation, mais les informations contenues dans l'espace tangent à une variété statistique sont peu exploitées. En géométrie riemannienne, les métriques dites naturelles sur les fibrés tangent donnent une vision de la géométrie de ces objets. Elles sont adaptées à la connexion de Levi-Civita qui n'est pas d'un usage courant en géométrie de l'information. Le travail présenté ici vise à définir un équivalent des métriques naturelles dans le cas de paires de connexions duales, qui ont une interprétation statistique.

Mots-clés. Géométrie de l'information, métriques naturelles, connexions duales, variété statistique, information de Fischer.

Abstract Information geometry provides an elegant framework for formulating statistical properties. Numerous works have been dedicated to the study of the relations between geometric representation and estimation, but the information contained in the tangent space to a statistical variety is only marginally considered. In Riemannian geometry, the so-called natural metrics on the tangent bundle give clues about the geometry of these objects. They are adapted to the Levi-Civita connection which is not of common use in information geometry. The work presented here aims to define an equivalent of natural metrics in the case of dual connections, for which a statistical interpretation exists.

Keywords. information geometry, natural metrics, dual connections, statistical manifold, Fischer information.

1. Introduction

La géométrie de l'information représente une famille de modèles statistiques comme des points sur une variété différentielle [1]. La métrique d'information de Fischer munit une telle variété d'une structure riemanienne. De nombreuses propriétés statistiques d'estimateurs peuvent se formuler directement en termes géométriques dans ce cadre. Il est ainsi possible de donner une interprétation élégante de résultats classiques, et dans certains cas d'en obtenir de nouveaux. Néanmoins, l'étude systématique de la géométrie des variétés statistiques est encore en développement, l'exploitation directe de résultats issus de la géométrie différentielle dans le cadre de la statistique étant assez limitée. Dans le travail présenté ici, une exploration d'un objet fondamental en géométrie de l'information, le fibré tangent à une variété statistique, sera abordé du point de vue du géomètre afin d'en déduire une interprétation nouvelle de la dualité entre connexion “mélange” et “exponentielle”. Pour ce faire, une métrique naturelle sera définie sur le fibré tangent à une variété statistique qui généralise les métriques naturelles associées à la connexion de Levi-Civita.

2. Connexions

De façon intuitive, une connexion donne un procédé pour dériver un champ de vecteurs dans une direction. Formellement, si $E \xrightarrow{\tau} M$ est un fibré vectoriel sur une variété M, une connexion ∇ sur E est une section ∇ du fibré $E \times TM^* \to M$ qui est une dérivation sur E :

$$\nabla_Z(\phi X + Y) = Z(\phi)X + \nabla_Z X + \nabla_Z Y \tag{1}$$

où $\phi\colon M \to \mathbb{R}$ est une application lisse.

Dans le cadre de la géométrie de l'information, on se placera presque toujours dans le cas $E = TM$.

Une connexion ∇ est dite sans torsion si pour tout couple de champs de vecteurs X, Y, on a :

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{2}$$

Enfin, si (M, g) est une variété riemanienne, la connexion ∇ sera dite métrique si :

$$Z(g(X, Y)) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) \tag{3}$$

L'unique connexion métrique et sans torsion sur une variété riemanienne (M, g) est appelée connexion de Levi-Civita.

Dans le cadre de la géométrie de l'information, la connexion de Levi-Civita n'est pas la plus intéressante à étudier. On relaxe alors la condition (3) en introduisant les connexions duales. Un couple ∇, ∇^* de connexions duales vérifie :

$$Z(g(X, Y)) = g(\nabla_Z X, Y) + g(X, \nabla^*_Z Y) \tag{4}$$

En utilisant les formes différentielles à valeurs vectorielles [3], on peut reformuler (3) comme :

$$dg(X, Y)(Z) + g(dX.Z, Y) + g(X, dY.Z) = g(\nabla_Z, X, Y) + g(X, \nabla^*_Z Y) \tag{5}$$

g étant non dégénérée, il existe pour toute forme linéaire $L \in TM^*$ une unique section $\tilde{L}$ de TM telle que pour tout $X \in TM$, $LX = g(X, \tilde{L})$. On en déduit donc l'existence de L_Z (resp. U_Z) sections de TM telles que :

$$dg(X, Y)(Z) = g(L_Z X, Y) \tag{6}$$

$$dg(X, Y)(Z) = g(X, U_Z Y) \tag{7}$$

g étant symétrique, on a $L_Z = U_Z$, et pour tout $\alpha \in [0, 1]$ on a :

(8)

$$(1 - \alpha)g(L_Z X, Y) + \alpha g(X, L_Z Y) + g(dX.Z, Y) + g(X, dY.Z) = g(\nabla_Z X, Y) + g(X, \nabla^*_Z Y)$$

L'égalité devant avoir lieu pour tout triplet (X, Y, Z) de sections, on obtient finalement :

$$\nabla_Z X = (1-\alpha) L_Z X + dX.Z \tag{9}$$

$$\nabla^*_Z Y = \alpha L_Z Y + dY.Z \tag{10}$$

Ces expressions sont précisément celles obtenues dans [1] sous le terme de α (resp. $-\alpha$) connexions, après application de la transformation $\alpha \mapsto 2\alpha - 1$. En coordonnées locales, on obtient facilement la relation :

$$g(L_{\partial_k}\partial_i, \partial_j) = E[\partial_i l \partial_j l \partial_k l] \tag{11}$$

avec l fonction de vraisemblance.

Le choix $\alpha = 1/2$ conduit à la métrique de Levi-Civita pour laquelle $\nabla = \nabla^*$. Les choix $\alpha = 0$ et $\alpha = 1$ correspondent respectivement aux connexions “mélange” et “exponentielles”.

3. Métriques naturelles sur les fibrés

Soit une variété M munie d'une métrique riemannienne g et soit $TM \overset{\pi}{\mapsto} M$ son fibré tangent. TM possédant également une structure de variété, il est naturel de rechercher des métriques dérivées de g qui puissent y être appliquées. Cette question a été traitée en particulier dans Sasaki [4] et Cheeger,Gromoll [2]. Le principe général est de décomposer l'espace tangent à TM au point (m, v) sous la forme d'une somme directe d'un sous-espace vertical et d'un sous-espace horizontal : $T_{(m,v)} = T^h_{(m,v)} \oplus T^v_{(m,v)}$. Le sous-espace vertical en (m, v) s'obtient comme le noyau en (m, v) de l'application $d\pi \colon TTM \mapsto TM$. L'espace horizontal $T^h_{(m,v)}$ n'est pas unique, mais on peut l'associer à une connexion ∇ de la façon suivante. Soit une courbe lisse $t \mapsto X(t) = (\gamma(t), U(t))$ de TM. Sa dérivée en 0 définit un vecteur tangent $X'(0)$ de TTM. On dira que la courbe $(\gamma(t), U(t))$ est horizontale si U est obtenu par transport parallèle selon ∇ le long de γ. On a alors $\nabla_{\gamma'(0)} U(0) = 0$. Un vecteur X' de TTM sera alors dit horizontal s'il est tangent à une courbe horizontale au sens précédent. On définit alors :

$$\tilde{\nabla}_{p,u} \colon TTM \to TM \tag{12}$$

l'application qui à un vecteur tangent X' associe $\nabla_{\gamma'(0)} U(0)$, avec $t \mapsto X(t) = (\gamma(t), U(t))$ courbe lisse de TM vérifiant $\gamma(0) = p$, $\gamma'(0) = u$. L'espace horizontal $T^h_{(m,v)}$ s'identifie alors au noyau $\ker \tilde{\nabla}_{(p,u)}$.

La construction précédente s'applique à toute connexion. Néanmoins, lorsque ∇ est la connexion de Levi-Civita, l'application $\tilde{\nabla}_{p,u}$ peut s'interpréter comme une projection orthogonale. Ceci permet de définir de façon intuitive des métriques dites naturelles sur TM. Si l'on considère deux champs X', Y' sur TTM se décomposant en composantes horizontales et verticales sous la forme $X' = X^h + X^v, Y' = Y^h + Y^v$ et relevant les champs X, Y sur TM, une métrique $\tilde{g}$ sur TTM sera dite naturelle si elle vérifie :

$$\tilde{g}(X^h, Y^h)_{(p,u)} = g_p(X, Y) \tag{13}$$

$$\tilde{g}(X^h, Y^v)_{(p,u)} = 0 \tag{14}$$

On notera qu'il s'agit d'une traduction directe de la propriété d'orthogonalité.

La métrique naturelle de Sasaki applique g aux composantes verticales $\tilde{g}_{(p,u)}(X^y, Y^v) = g_p(X, Y)$. Elle impose des conditions de rigidité très fortes. La métrique de Cheeger-Gromoll s'exprime sur les composantes verticales par :

$$\tilde{g}(X^v, Y^v)_{(p,u)} = \frac{1}{1 + g(u,u)} \left(g_p(X, Y) + g_p(X, u) g_p(Y, u) \right) \tag{15}$$

Dans le cas du fibré associé à une variété statistique, la connexion de Levi-Civita n'est qu'une des α-connexions et ne présente pas un intérêt particulier. Il convient donc de généraliser la construction précédente aux couples de connexions duales.

4. Connexions naturelles duales

L'idée sous-jacente est de partir de l'interprétation d'une connexion ∇ comme une projection de TTM sur TM. Soit P une matrice de projection. En utilisant la décomposition en valeurs singulière, on peut l'écrire sous une forme factorisée $P = UV^*$ où les matrices sont des isométries partielles. Un projecteur orthogonal admet quant à lui une forme $P = UU^*$. Dans le cas des connexions, on peut établir une décomposition de même nature : $\nabla = UV^*$ où V^* traduit un relèvement de TM sur l'espace horizontal de TTM et U une opération inverse (submersion de TTM sur TM). Si ∇, ∇^* est un couple de connexions duales, alors $\nabla^* = VU^*$ au sens précédent.

Si l'on se donne une métrique g sur la variété de départ M, on peut étendre la notion de métrique naturelle sur TTM, en partant de la décomposition $\nabla = UV^*$ en imposant :

$$\tilde{g}(U^*X, V^*Y) = g(X, Y) \tag{16}$$

$$\tilde{g}(X^v, (U^* + V^*)Y) = 0 \tag{17}$$

En raison de la symétrie de g, la condition 16 s'écrit également :

$$\tilde{g}(V^*X, U^*Y) = g(X, Y) \tag{18}$$

et en utilisant la bilinéarité de $\tilde{g}$:

$$\frac{1}{2} \left(\tilde{g}((U^* + V^*)X, (U^* + V^*)Y) - \tilde{g}(U^*X, U^*y) - \tilde{g}(V^*X, V^*Y) \right) = g(X, Y) \tag{19}$$

On remarquera que l'on retrouve la notion de métrique naturelle dans le cas de la connexion de Levi-Civita. Enfin, on notera que des connexions duales induisent la même métrique horizontale.

Sur les variétés statistiques, les α-connexions forment naturellement des systèmes duaux. Le procédé précédent s'applique et donne accès à des métriques naturelles sur le fibré tangent. La métrique de Sasaki est la plus simple à étendre. Elle se traduit dans l'espace vertical par une information de Fischer au second ordre, à savoir faisant intervenir des produits de dérivées secondes de la log-vraisemblance.

Références

[1] S. Amari and H. Nagaoka. *Methods of Information Geometry.* Translations of mathematical monographs. American Mathematical Society, 2007.

[2] Jeff Cheeger and Detlef Gromoll. On the structure of complete manifolds of nonnegative curvature. *Annals of Mathematics*, 96(3) :413–443, 1972.

[3] D. Husemöller. *Fibre Bundles.* Graduate Texts in Mathematics. Springer, 1994.

[4] Shigeo Sasaki. On the differential geometry of tangent bundles of riemannian manifolds. *Tohoku Math. J. (2)*, 10(3) :338–354, 1958.