HYPERSURFACES IN STATISTICAL MANIFOLDS

FURUHATA Hitoshi

祝還暦 浦川肇 先生

古畑　仁
北海道大学大学院理学研究院

Let H be the $(n+1)$-dimensional upper half Hessian space of constant Hessian curvature 4. A horosphere $f_0 : \mathbb{R}^n \ni {}^t(y^1, \dots, y^n) \mapsto {}^t(y^1, \dots, y^n, b) \in H$ (b =constant> 0), as a hypersurface in a statistical manifold, induces a trivial Hessian structure (∇, g) on the source manifold $\mathbb{R}^n$. Conversely, we can characterize f_0 as such a hypersurface(Corollary 3.2). This is a joint work with KUROSE Takashi.

1. Statistical Manifolds

Throughout this article, M denotes an n-dimensional manifold, ∇ an affine connection on M, and g a Riemannian metric on M. We denote by $\Gamma(E)$ the set of sections of the vector bundle $E \to M$. For example, $\Gamma(TM^{(p,q)})$ means the set of tensor fields of type (p,q) on M. All the objects are assumed to be smooth, and treated in the local geometry.

Definition 1.1. A pair (∇, g) is called *statistical structure* on M if (1) ∇ is of torsion free, and (2) $\nabla g \in \Gamma(TM^{(0,3)})$ is totally symmetric.

We can naturally find statistical structures as follows. Let ∇^g be the the Levi-Civita connection of g. By definition, a pair (∇^g, g) is a statistical structure, which may be called a *trivial* statistical structure. Let ∇^* be the connection defined by

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z)$$

for any $X, Y, Z \in \Gamma(TM)$, which is called the *dual connection* of ∇ with respect to g. If (∇, g) is a statistical structure on M, so is (∇^*, g).

Let $(\widetilde{M}, \widetilde{\nabla}, \widetilde{g})$ be a statistical manifold, and $f : M \to \widetilde{M}$ an immersion. We define g and ∇ on M by

$$g := f^*\widetilde{g}, \quad g(\nabla_X Y, Z) = \widetilde{g}(\widetilde{\nabla}_X f_* Y, f_* Z) \tag{1.1}$$

for any $X, Y, Z \in \Gamma(TM)$, where the connection induced from $\widetilde{\nabla}$ by f on the induced bundle $f^*T\widetilde{M} \to M$ is denoted by the same symbol $\widetilde{\nabla}$. Then the pair (∇, g) is a statistical structure on M, which is called the one *induced by f from* $(\widetilde{\nabla}, \widetilde{g})$.

The name of this structure comes from information geometry (See [1]). Let $p(\cdot, \theta) : (X, dx) \to (0, \infty)$ be a probability density parametrized by $\theta = (\theta^1, \dots, \theta^n) \in \Theta \subset \mathbb{R}^n$. For any constant $\alpha \in \mathbb{R}$, we set

$$g_\theta := \sum \{ \int_X \frac{\partial \log p}{\partial \theta^i}(x, \theta) \frac{\partial \log p}{\partial \theta^j}(x, \theta) p(x, \theta) dx \} d\theta^i d\theta^j,$$

and

$$\Gamma^{(\alpha)}_{ijk}(\theta) := \int_X \{ \frac{\partial^2 \log p}{\partial \theta^i \partial \theta^j}(x, \theta) + \frac{1-\alpha}{2} \frac{\partial \log p}{\partial \theta^i}(x, \theta) \frac{\partial \log p}{\partial \theta^j}(x, \theta) \} \frac{\partial \log p}{\partial \theta^k}(x, \theta) p(x, \theta) dx.$$

It is easy to see that g_θ is a positive semi-definite quadratic form on $T_\theta \Theta$. If g is a Riemannian metric on Θ, then $(\Theta, \nabla^{(\alpha)}, g)$ is a statistical manifold, where $\nabla^{(\alpha)}$ is a connection defined by $\Gamma^{(\alpha)}_{ijk} = g(\nabla^{(\alpha)}_{\frac{\partial}{\partial \theta^i}} \frac{\partial}{\partial \theta^j}, \frac{\partial}{\partial \theta^k})$. Then g is called the *Fisher metric* and $\nabla^{(\alpha)}$ the Amari's α-connection with respect to $\{p(\cdot, \theta) \mid \theta \in \Theta\}$.

For example, a family $\{p(\cdot, \theta) \mid \theta \in \Theta\}$ is given as follows. The normal distribution with mean μ and variance σ^2 is written as

$$N(x, \mu, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2}(x-\mu)^2 \right\}, \quad x \in \mathbb{R},\ \mu \in \mathbb{R},\ \sigma > 0.$$

Set $\Theta := \{(\theta^1, \theta^2) \in \mathbb{R}^2 \mid \theta_2 > 0\}$, $X := \mathbb{R}$ and $p(x, \theta) := N(x, \sqrt{2}\theta^1, (\theta^2)^2)$. Then the statistical manifold with respect to $\{p(\cdot, \theta) \mid \theta \in \Theta\}$ has a Riemannian metric $g = 2(\theta^2)^{-2} \sum d\theta^i d\theta^i$ of constant curvature $-1/2$ and a flat connection $\nabla^{(-1)}$ with

$$\nabla^{(-1)}_{\frac{\partial}{\partial \theta^1}} \frac{\partial}{\partial \theta^1} = 2(\theta^2)^{-1} \frac{\partial}{\partial \theta^2}, \quad \nabla^{(-1)}_{\frac{\partial}{\partial \theta^2}} \frac{\partial}{\partial \theta^2} = (\theta^2)^{-1} \frac{\partial}{\partial \theta^2},$$
$$\nabla^{(-1)}_{\frac{\partial}{\partial \theta^1}} \frac{\partial}{\partial \theta^2} = \nabla^{(-1)}_{\frac{\partial}{\partial \theta^2}} \frac{\partial}{\partial \theta^1} = 0.$$

Definition 1.2. A statistical structure (∇, g) is said to be *Hessian* if ∇ is flat. For a statistical structure (∇, g) we define the *difference tensor* $K^{(\nabla,g)} \in \Gamma(TM^{(1,2)})$ as

$$K^{(\nabla,g)}(X,Y) := \nabla_X Y - \nabla^g_X Y.$$

A Hessian structure (∇, g) is *of constant Hessian curvature* $c \in \mathbb{R}$ if

$$\left(\nabla_X K^{(\nabla,g)}\right)(Y,Z) = -\frac{c}{2}\left\{g(X,Y)Z + g(X,Z)Y\right\}$$

for any $X, Y, Z \in \Gamma(TM)$.

Concerning Hessian geometry, we refer the reader to [3]. The tangent bundle of a Hessian manifold of constant Hessian curvature c has the Kählerian structure of constant holomorphic sectional curvature $-c$. It is also known that the Riemannian metric of a Hessian manifold of constant Hessian curvature c has constant sectional curvature $-c/4$.

Example 1.3. Generalizing the above statistical manifold of normal distributions, we construct a Hessian manifold of positive constant Hessian curvature as follows. Let $(H, \widetilde{g})$ be the upper half space of constant curvature -1:

$$\begin{aligned} &H := \left\{y = {}^t(y^1, \dots, y^{n+1}) \in \mathbb{R}^{n+1} \mid y^{n+1} > 0\right\}, \\ &\widetilde{g} := (y^{n+1})^{-2} \textstyle\sum_{A=1}^{n+1} dy^A dy^A. \end{aligned}$$

We define the affine connection $\widetilde{\nabla}$ on H by the following relations:

$$\begin{gathered} \widetilde{\nabla}_{\frac{\partial}{\partial y^{n+1}}} \frac{\partial}{\partial y^{n+1}} = (y^{n+1})^{-1} \frac{\partial}{\partial y^{n+1}}, \quad \widetilde{\nabla}_{\frac{\partial}{\partial y^i}} \frac{\partial}{\partial y^j} = 2\delta_{ij}(y^{n+1})^{-1} \frac{\partial}{\partial y^{n+1}}, \\ \widetilde{\nabla}_{\frac{\partial}{\partial y^i}} \frac{\partial}{\partial y^{n+1}} = \widetilde{\nabla}_{\frac{\partial}{\partial y^{n+1}}} \frac{\partial}{\partial y^j} = 0, \end{gathered}$$

where $i, j = 1, \dots, n$. Then $(H, \widetilde{\nabla}, \widetilde{g})$ is a Hessian manifold of constant Hessian curvature 4. Moreover, we conjecture that $\widetilde{\nabla}$ is the only connection of constant Hessian curvature 4 globally defined on $(H, \widetilde{g})$.

2. Statistical Hypersurfaces

Definition 2.1. Let (M, ∇, g) and $(\widetilde{M}, \widetilde{\nabla}, \widetilde{g})$ be two statistical manifolds. An immersion $f : M \to \widetilde{M}$ is called a *statistical* immersion if (1.1) holds.

This notion seems useful in statistical inference. Concerning the theory on statistical submanifolds, we refer the reader to [4].

Example 2.2. Let $(H, \widetilde{\nabla}, \widetilde{g})$ be the $(n+1)$-dimensional upper half Hessian space of constant Hessian curvature 4 as in Example 1.3. For a constant $b > 0$, denote the following immersion by f_0:

$$\mathbb{R}^n \ni {}^t(y^1, \dots, y^n) \mapsto {}^t(y^1, \dots, y^n, b) \in H.$$

Let (∇, g) be the statistical structure on $\mathbb{R}^n$ induced by f_0 from $(\widetilde{\nabla}, \widetilde{g})$. We then get that (∇, g) is a Hessian structure and $K^{(\nabla,g)} = 0$. In other words, f_0 is a statistical immersion of the trivial Hessian manifold $(\mathbb{R}^n, \nabla, g)$ into the upper half Hessian space $(H, \widetilde{\nabla}, \widetilde{g})$.

The goal of this article is to prove that the above f_0 is the only statistical immersion of a trivial Hessian manifold into $(H, \widetilde{\nabla}, \widetilde{g})$. It should be remarked that in Riemannian geometry we can not characterize f_0 as a flat hypersurface. In fact, there are many flat surfaces in the upper half space (See [2], for example). To state our theorem more precisely, we fix the notation as follows. Let $f : (M, \nabla, g) \to (H, \widetilde{\nabla}, \widetilde{g})$ be a statistical immersion, and $\xi \in \Gamma(f^*TH)$ a unit normal vector field of f. We define $h, h^* \in \Gamma(TM^{(0,2)})$, $A, A^* \in \Gamma(TM^{(1,1)})$ and $\tau, \tau^* \in \Gamma(TM^*)$ by the following Gauss and Weingarten formulas:

$$\begin{aligned}
&\widetilde{\nabla}_X f_* Y = f_* \nabla_X Y + h(X,Y)\xi, \\
&\widetilde{\nabla}_X \xi = -f_* A^* X + \tau^*(X)\xi, \\
&\widetilde{\nabla}^*_X f_* Y = f_* \nabla^*_X Y + h^*(X,Y)\xi, \\
&\widetilde{\nabla}^*_X \xi = -f_* AX + \tau(X)\xi, \quad X, Y \in \Gamma(TM),
\end{aligned} \tag{2.1}$$

where $\widetilde{\nabla}^*$ is the dual connection of $\widetilde{\nabla}$ with respect to $\widetilde{g}$. It is easy to show that the connection induced from $\widetilde{\nabla}^*$ is the dual connection of ∇ with respect to g. Besides, the following hold for any $X, Y \in \Gamma(TM)$.

$$\begin{aligned}
&h(X,Y) = g(AX,Y), \quad h^*(X,Y) = g(A^*X,Y), \\
&\tau(X) + \tau^*(X) = 0.
\end{aligned} \tag{2.2}$$

In addition, we define $II \in \Gamma(TM^{(0,2)})$ and $S \in \Gamma(TM^{(1,1)})$ by using the Riemannian Gauss and Weingarten formulas:

$$\begin{aligned}
&\nabla^{\widetilde{g}}_X f_* Y = f_* \nabla^g_X Y + II(X,Y)\xi, \\
&\nabla^{\widetilde{g}}_X \xi = -f_* SX.
\end{aligned} \tag{2.3}$$

Proposition 2.3. *Let $f_0 : \mathbb{R}^n \to H$ be the statistical hypersurface as in Example 2.2. Then we calculate the above quantities as below:*

$$\begin{aligned}
&\xi = b\frac{\partial}{\partial y^{n+1}},\\
&II = g, \quad S = I,\\
&h = 2g,\ h^* = 0,\ A^* = 0,\ A = 2I,\ \tau^* = \tau = 0.
\end{aligned}$$

Example 2.4. We denote by f_1 the other type expression of horospheres in the upper half space, that is,

$$f_1 : \mathbb{R}^n \ni z \mapsto \begin{bmatrix} 4r^2(|z-a|^2+4r^2)^{-1}(z-a)+a \\ 8r^3(|z-a|^2+4r^2)^{-1} \end{bmatrix} \in H,$$

where $r > 0$ and $a \in \mathbb{R}^n$. We remark that the image $f_1(\mathbb{R}^n)$ is the set $\{y \in \mathbb{R}^{n+1} \mid |y - \begin{bmatrix} a \\ r \end{bmatrix}| = r\} \setminus \{\begin{bmatrix} a \\ 0 \end{bmatrix}\}$, which is congruent to $f_0(\mathbb{R}^n)$ in the sense of Riemannian geometry. We can take $\xi = 8r^3(|z-a|^2 + 4r^2)^{-2}\left\{4r\sum(z^i-a^i)\frac{\partial}{\partial y^i} + (4r^2 - |z-a|^2)\frac{\partial}{\partial y^{n+1}}\right\}$ as a unit normal vector field, and calculate that $II = g$, $S = I$ and ∇^g is flat, and that $\tau^* = -r^{-2}(|z-a|^2+4r^2)^{-3}\left(|z-a|^6 + 6|z-a|^4r^2 - 32r^6\right)\sum(z^i - a^i)dz^i \neq 0$.

3. Results

We are in the position to state our theorem, which assures that a statistical hypersurface with trivial Hessian structure and a horosphere in the upper half Hessian space are congruent in the sense of Riemannian geometry.

Theorem 3.1. *If $f : (M, \nabla, g) \to (H, \widetilde{\nabla}, \widetilde{g})$ is a statistical immersion of a trivial Hessian manifold, then $S = \pm I$. Moreover, $A^* = 0$ and $\tau^* = 0$ hold.*

Combining it with Example 2.4, we can prove

Corollary 3.2. *Let (M, ∇, g) be a connected trivial Hessian manifold. If $f : (M, \nabla, g) \to (H, \widetilde{\nabla}, \widetilde{g})$ is a statistical immersion, then $f(M)$ is an open subset of $f_0(\mathbb{R}^n)$.*

In order to get Theorem 3.1, we need a series of lemmas obtained by using the Gauss, Codazzi and Ricci equations. Though the proof is omitted, we place one of them at the last of this article for your information.

Lemma 3.3. *Let $f : (M, \nabla, g) \to (H, \widetilde{\nabla}, \widetilde{g})$ be a statistical hypersurface in the upper half Hessian space. If $K^{(\nabla,g)} = 0$, then $\tau^* = 0$.*

References

[1] Amari, S. and Nagaoka, H., *Methods of Information Geometry*, American Mathematical Society, 2000.
[2] Galvez, J.A., Martinez, A. and Milan, F., Flat surfaces in the hyperbolic 3-space, Math. Ann., **316**(2000), 419–435.
[3] Shima, H., *Hessian Geometry*, Shokabo, 2001 (in Japanese).
[4] Vos, P.W., Fundamental equations for statistical submanifolds with applications to the Bartlett correction, Ann. Inst. Statist. Math. **41**(1989), 429–450.

Department of Mathematics, Hokkaido University, Sapporo 060-0810, Japan
E-mail address: `furuhata@math.sci.hokudai.ac.jp`