



Title	Local existence of statistical diffeomorphisms
Author(s)	Satoh, Naoto
Citation	JP Journal of Geometry and Topology, 22(1), 73-95 https://doi.org/10.17654/GT022010073
Issue Date	2019-02
Doc URL	https://hdl.handle.net/2115/75440
Type	journal article
File Information	JP JGT22-1_73-95.pdf





LOCAL EXISTENCE OF STATISTICAL DIFFEOMORPHISMS

Naoto Satoh

Department of Mathematics

Graduate School of Science

Hokkaido University

Japan

Abstract

A diffeomorphism between statistical manifolds is said to be statistical if it preserves statistical structures. Our purpose is to find conditions that guarantee an extension of a given linear isomorphism between given tangent spaces to a local statistical diffeomorphism. In Riemannian geometry, it is known as the Cartan-Ambrose-Hicks theorem, which implies that a Riemannian metric is locally determined by its Riemannian curvature tensor. We generalize this theorem for statistical manifolds, and, in particular, for Hessian manifolds. We prove that a statistical structure is locally characterized by its Riemannian curvature tensor and difference tensor. Furthermore, we show that a Hessian structure is locally determined by its Hessian curvature tensor and difference tensor.

1. Introduction

Information geometry has begun from the study of geometrical properties of a smooth family of parametrized probability distributions,

Received: August 11, 2018; Accepted: September 22, 2018

2010 Mathematics Subject Classification: 53A15, 53B20, 53B99.

Keywords and phrases: information geometry, statistical diffeomorphism, statistical manifold, Hessian manifold.

which we call a statistical model. A statistical model is regarded as a manifold, and naturally equipped with not only a Riemannian metric but also a pair of dual connections. Information geometry aims to study statistical models from the viewpoint of dual connections, and provides new methods for so many fields, for example, statistics, information theory, and machine learning, where probability distributions play crucial roles (see [1, 2, 4]). Statistical manifold is a geometric generalization of statistical model.

Definition 1.1. Let M be a manifold endowed with a torsion-free connection ∇ and a Riemannian metric g . We say that (M, ∇, g) is a *statistical manifold* if the $(0, 3)$ -tensor field $C := \nabla g$, which we call the cubic form, is totally symmetric. The pair (∇, g) is said to be the *statistical structure* on M .

Then we can define another torsion-free connection ∇^* by

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

for any vector fields $X, Y, Z \in \Gamma(TM)$. From this definition, it is easy to see that (∇^*, g) is also a statistical structure on M and $(\nabla^*)^* = \nabla$. For this reason, we call ∇^* the *dual connection* of ∇ with respect to g .

We remark that a statistical manifold with $C = 0$ is a Riemannian manifold with the Levi-Civita connection of g , denoted by ∇^g . In this sense, statistical manifold can be viewed as a generalization of Riemannian manifold.

In Riemannian geometry, isometry, which is a metric preserving diffeomorphism between Riemannian manifolds, plays important roles. Correspondingly, in geometry of statistical manifolds, it is meaningful that we consider statistical structure preserving diffeomorphisms.

Definition 1.2. Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two statistical manifolds. Then a diffeomorphism $\varphi : M \rightarrow \bar{M}$ is called a *statistical*

diffeomorphism if

$$\varphi_*(\nabla_X Y) = \bar{\nabla}_{\varphi_* X} \varphi_* Y, \quad g(X, Y) = \bar{g}(\varphi_* X, \varphi_* Y)$$

for any vector fields $X, Y \in \Gamma(TM)$. Then (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ are said to be *statistical diffeomorphic*.

It is a natural question to ask which conditions guarantee an extension of a given linear isomorphism to a statistical diffeomorphism. In Riemannian geometry, this question was treated by Cartan [5], and expanded by Ambrose [3] and Hicks [9]. The Cartan-Ambrose-Hicks theorem in a local setting is stated as follows:

Theorem 1.3. *Let (M, g) and $(\bar{M}, \bar{g})$ be two Riemannian manifolds of the same dimension, and $p \in M$ and $\bar{p} \in \bar{M}$. Let $\Phi : T_p M \rightarrow T_{\bar{p}} \bar{M}$ be a linear isometry. Set φ and Φ_γ as (3.1) and (3.2) (the details shall be mentioned in Setting 3.1). Suppose that for any ∇^g -geodesic γ starting at p we have*

$$\Phi_\gamma(R^g(u, v)w) = R^{\bar{g}}(\Phi_\gamma(u), \Phi_\gamma(v))\Phi_\gamma(w)$$

for any $u, v, w \in T_q M$, where R^g and $R^{\bar{g}}$ are their Riemannian curvature tensor fields, respectively. Then φ is a local isometry around p and $\bar{p}$.

This theorem means, in a sense, that a Riemannian metric is locally determined by its Riemannian curvature tensor field, and it is a key fact implying that a space of constant Riemannian curvature is rich in isometries.

We shall generalize this theorem for statistical manifolds (Theorem 3.2). This theorem asserts that a statistical structure is locally characterized by its Riemannian curvature tensor field R^g and difference tensor field $K := \nabla - \nabla^g$ and their behavior under ∇^g -parallel translation along ∇^g -geodesics. In the case where $K = 0$, this theorem coincides with the Cartan-Ambrose-Hicks theorem above.

In particular, we consider the case of Hessian manifolds, which is an important subclass of statistical manifolds.

Definition 1.4. A statistical manifold (M, ∇, g) is said to be a *Hessian manifold* if ∇ is flat. The pair (∇, g) is called the *Hessian structure* of M .

For each point p of a Hessian manifold (M, ∇, g) , there exist a neighborhood U and a function $\psi \in C^\infty(U)$ such that $g|_U = \nabla d\psi$, that is, g

can be locally expressed by $[g_{ij}]_{i,j=1,\dots,n} = \left[\frac{\partial^2 \psi}{\partial x^i \partial x^j} \right]_{i,j=1,\dots,n}$ ($n = \dim M$),

where $\{x^1, \dots, x^n\}$ is an affine coordinate system with respect to ∇ . It is known that exponential families, for example, normal distributions and multinomial distributions, carry Hessian structures (see [15]). Hence, the study of Hessian manifolds is of great interest in information geometry.

We shall show that a Hessian structure is locally determined by its Hessian curvature tensor field $Q := \nabla K$ and difference tensor field K and their behavior under ∇^g -parallel translation along ∇^g -geodesics (Theorem 4.1).

This paper consists of four sections. In Section 2, we shall review preliminary facts on statistical manifolds, geodesics, exponential maps, and isometries. In Section 3, we shall discuss statistical diffeomorphisms and prove Theorem 3.2. Section 4 treats Hessian manifolds and is devoted to proving Theorem 4.1. In addition, we explain examples of statistical diffeomorphisms.

2. Preliminaries

Throughout this paper, let M be an n -dimensional manifold, ∇ be a torsion-free affine connection and g be a Riemannian metric on M . Assume that all the objects are smooth. We denote by $\Gamma(E)$ the set of sections of a

vector bundle $\pi : E \rightarrow M$. For example, $\Gamma(TM^{(p,q)})$ means the set of tensor fields of type (p, q) on M .

2.1. Statistical manifolds and Hessian manifolds

Let (M, ∇, g) be a statistical manifold. Now, in addition to the given connection ∇ , we may consider the Levi-Civita connection ∇^g for the given Riemannian metric g . Then we consider the difference tensor $K \in \Gamma(TM^{(1,2)})$ as follows:

Definition 2.1. For a statistical structure (∇, g) , we define the *difference tensor field* $K := K^{(\nabla, g)} \in \Gamma(TM^{(1,2)})$ as

$$K(X, Y) := \nabla_X Y - \nabla_X^g Y$$

for any $X, Y \in \Gamma(TM)$. We often denote $K(X, Y)$ by $K_X Y$ as well.

It is well-known that the difference tensor K holds the following properties:

$$K(X, Y) = K(Y, X),$$

$$g(K(X, Y), Z) = g(Y, K(X, Z)),$$

$$K(X, Y) = \nabla_X^g Y - \nabla_X^* Y = \frac{1}{2}(\nabla_X Y - \nabla_X^* Y).$$

We remark that a statistical structure (∇, g) is Riemannian if the difference tensor K vanishes. We can now relate the difference tensor K to the cubic form $C := \nabla g$ as follows:

$$C(X, Y, Z) = -2g(K(X, Y), Z).$$

For an affine connection ∇ , we set

$$R^\nabla(X, Y)Z := \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z$$

for $X, Y, Z \in \Gamma(TM)$. Then $R^\nabla \in \Gamma(TM^{(1,3)})$ is called the *curvature*

tensor field of ∇ . In particular, for the case of ∇^g , R^{∇^g} is said to be the Riemannian curvature tensor. From now on, we denote R^∇ by R , R^{∇^*} by R^* , and R^{∇^g} by R^g , respectively, for simplicity. We say that (M, g) has a Riemannian structure of constant Riemannian curvature c if there exists $c \in \mathbb{R}$ such that

$$R^g(X, Y)Z = c\{g(Y, Z)W - g(X, Z)Y\}$$

for any $X, Y, Z \in \Gamma(TM)$. The following formulas are obtained by direct calculation.

Lemma 2.2. *For a statistical manifold (M, ∇, g) , the following hold for $X, Y, Z \in \Gamma(TM)$:*

$$R^g(X, Y)Z = R(X, Y)Z - (\nabla_X K)(Y, Z) + (\nabla_Y K)(X, Z) + [K_X, K_Y]Z, \quad (2.1)$$

$$g(R(X, Y)Z, W) = -g(R^*(X, Y)W, Z). \quad (2.2)$$

From (2.2), we find that if R is vanished, then so is R^* . A torsion-free affine connection ∇ is said to be *flat* if its curvature tensor R is identically 0. Assume that (∇, g) is a Hessian structure, and connections ∇ and ∇^* are both flat. For this reason, a Hessian structure is also called a *dually flat structure* in information geometry.

Next, we define the Hessian curvature tensor $Q \in \Gamma(TM^{(1,3)})$ on a Hessian manifold (M, ∇, g) as follows:

Definition 2.3 [15]. (1) For a Hessian structure (∇, g) , we define the Hessian curvature tensor $Q \in \Gamma(TM^{(1,3)})$ by

$$Q(X, Y)Z := -(\nabla_X K)(Y, Z) = -(\nabla K)(Y, Z; X)$$

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

(2) A Hessian manifold (M, ∇, g) is said to be of *constant Hessian sectional curvature* $c \in \mathbb{R}$ if

$$Q(X, Y)Z = -\frac{c}{2} \{g(X, Y)Z + g(X, Z)Y\}$$

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

The Hessian curvature tensor Q plays a similar role to that of the Riemannian curvature tensor for a Kählerian metric. In fact, we can construct a Kählerian metric g^{TM} on the tangent bundle TM over a Hessian manifold (M, ∇, g) by the diagonal lifting (or the Sasakian lifting), and (TM, g^{TM}) is of constant holomorphic sectional curvature $-c$ if and only if (M, ∇, g) is of constant Hessian sectional curvature c .

From the following proposition, it can be seen that the Hessian curvature tensor Q carries more detailed information than the Riemannian curvature R^g .

Proposition 2.4 [15]. *Let (M, ∇, g) be a Hessian manifold and R^g be the Riemannian curvature tensor for g . Then the following holds:*

$$R^g(X, Y)Z = -\frac{1}{2} \{Q(X, Y)Z - Q(Y, X)Z\}.$$

Furthermore, if (M, ∇, g) is of constant Hessian curvature c , then (M, ∇, g) has a Riemannian structure of constant Riemannian curvature $-c/4$.

2.2. Geodesics and exponential maps

Let M be a smooth manifold with a torsion-free affine connection ∇ , and let $\gamma : I \rightarrow M$ be a smooth curve with parameter $t \in I \subset \mathbb{R}$. We denote by $\Gamma(\gamma^{-1}TM)$ the set of vector fields along γ on M . If the tangent vector fields $\dot{\gamma} \in \Gamma(\gamma^{-1}TM)$ are autoparallel relative to ∇ , that is,

$$\nabla_{\dot{\gamma}} \dot{\gamma} = 0, \quad (2.3)$$

then we say that $\gamma(t)$ is a ∇ -geodesic.

Let $(U; \phi = (u^1, \dots, u^n))$ be local coordinates for M and write $\gamma^i(t) := u^i(\gamma(t))$ ($i = 1, \dots, n$). Then we can rewrite (2.3) as follows:

$$\ddot{\gamma}^k(t) + \sum_{i,j=1}^n \Gamma_{ij}^k(\gamma(t)) \dot{\gamma}^i(t) \dot{\gamma}^j(t) = 0 \quad (k = 1, \dots, n),$$

where $\Gamma_{ij}^k \in C^\infty(U)$ ($i, j, k = 1, \dots, n$) are the Christoffel symbols for ∇ .

This constitutes a system of linear second order ODEs. Therefore, for given $p \in M$, $v \in T_p M$, there exists a unique geodesic γ_v with a maximal open interval $I_v \subset \mathbb{R}$ with $0 \in I_v$ such that $\gamma_v(0) = p$ and $\dot{\gamma}_v(0) = v$.

If we set for $p \in M$,

$$D_p := \{v \in T_p M; 1 \in I_v\} \subset T_p M,$$

then D_p is the star-shaped neighborhood of the origin $0_p \in T_p M$. We can then define the ∇ -exponential map Exp_p^∇ on D_p as follows:

Definition 2.5. We define a mapping $\text{Exp}_p^\nabla : D_p \rightarrow M$ by

$$\text{Exp}_p^\nabla v := \gamma_v(1) \quad (v \in D_p),$$

and call it the ∇ -exponential map at p .

We then have $\text{Exp}_p^\nabla(tv) = \gamma_v(t)$ for any $t \in I_v$. Now, if we identify $T_{0_p}(T_p M) = T_p M$, then the differential $(\text{Exp}_p^\nabla)_{*0_p}$ of Exp_p^∇ at 0_p turns out to be the identity map of $T_p M$. Thus, by the inverse function theorem, there exists a convex open neighborhood $\mathcal{U}_p \subset D_p$ of 0_p such that $\text{Exp}_p^\nabla : \mathcal{U}_p \rightarrow U := \text{Exp}_p^\nabla(\mathcal{U}_p)$ is a local diffeomorphism. If we take a

basis $\{e_1, \dots, e_n\}$ of $T_p M$ and consider the correspondence

$$\text{Exp}_p^\nabla \left(\sum_i u^i e_i \right) \mapsto (u^1, \dots, u^n) \quad \left(\sum_i u^i e_i \in \mathcal{U}_p \right),$$

then we obtain a local coordinate system around $p \in U$, which we call a ∇ -normal coordinate system.

2.3. Properties of isometries

Let (M, g) and $(\bar{M}, \bar{g})$ be two Riemannian manifolds and let $\varphi : M \rightarrow \bar{M}$ be a smooth map. Then a mapping φ is said to be a *local isometry* if for each $p \in M$ the differential $(\varphi_*)_p$ is a linear isometry relative to $g_p, \bar{g}_{\varphi(p)}$, that is, $g_p(u, v) = \bar{g}_{\varphi(p)}((\varphi_*)_p(u), (\varphi_*)_p(v))$ for any $u, v \in T_p M$. Furthermore, if φ is a diffeomorphism of M onto $\bar{M}$, then we say that φ is an *isometry*, and see

$$g(X, Y) = \bar{g}(\varphi_* X, \varphi_* Y)$$

for every $X, Y \in \Gamma(TM)$. In a local isometry case, it is to be noted that each point $p \in M$ has a neighborhood U such that $\varphi|_U$ is an isometry of U onto a neighborhood of $\varphi(p) \in \bar{M}$ because of the inverse function theorem.

Notice that an isometry $\varphi : (M, g) \rightarrow (\bar{M}, \bar{g})$ has the following properties:

$$(1) \quad \varphi_*(\nabla_X^g Y) = \nabla_{\varphi_* X}^{\bar{g}} \varphi_* Y \quad \text{for all } X, Y \in \Gamma(TM).$$

(2) Let P_c be the ∇^g -parallel translation from $c(a)$ to $c(b)$ along a smooth curve $c : [a, b] \rightarrow M$, and let $\bar{P}_{\varphi \circ c}$ be the $\nabla^{\bar{g}}$ -parallel translation from $\varphi \circ c(a)$ to $\varphi \circ c(b)$ along $\varphi \circ c$ in $\bar{M}$. Then $(\varphi_*)_{c(b)} \circ P_c = \bar{P}_{\varphi \circ c} \circ (\varphi_*)_{c(a)}$.

(3) If γ is a ∇^g -geodesic in M , then $\varphi \circ \gamma$ is a $\nabla^{\bar{g}}$ -geodesic in $\bar{M}$.

(4) Let Exp_p be the ∇^g -exponential map at $p \in M$ and $\overline{\text{Exp}}_{\varphi(p)}$ the $\nabla^{\bar{g}}$ -exponential map at $\varphi(p) \in \bar{M}$. Then $\varphi \circ \text{Exp}_p = \overline{\text{Exp}}_{\varphi(p)} \circ (\varphi_*)_p$, whenever the left-hand-side is defined.

(5) Since φ preserves both brackets and covariant derivatives, it also preserves Riemannian curvature tensors:

$$\varphi_*(R^g(X, Y)Z) = R^{\bar{g}}(\varphi_*X, \varphi_*Y)\varphi_*Z.$$

We see that a local isometry is uniquely determined by its differential map at a single point from the following proposition:

Proposition 2.6. *Let (M, g) and $(\bar{M}, \bar{g})$ be Riemannian manifolds, and $\varphi, \psi : M \rightarrow \bar{M}$ be local isometries. If there is a point $p \in M$ such that $(\varphi_*)_p = (\psi_*)_p$, then $\varphi = \psi$.*

Proof. Set $A = \{q \in M \mid (\varphi_*)_q = (\psi_*)_q\}$. By continuity, A is closed in M , and A is nonempty because A contains p . Thus, it suffices to show that A is open. We assert that if $q \in A$, then any ∇^g -normal neighborhood U of q is contained in A . If $q' \in U$, then there exists a vector $v \in T_q M$ such that $\gamma_v(1) = \text{Exp}_q(v) = q'$. Hence,

$$\begin{aligned} \varphi(q') &= \varphi(\gamma_v(1)) \\ &= (\overline{\text{Exp}}_{\varphi(q)} \circ (\varphi_*)_q \circ (\text{Exp}_q)^{-1})(\text{Exp}_q(v)) \\ &= \overline{\text{Exp}}_{\varphi(q)}((\varphi_*)_q(v)) \\ &= \bar{\gamma}_{(\varphi_*)_q(v)}(1) \\ &= \bar{\gamma}_{(\psi_*)_q(v)}(1) \end{aligned}$$

$$\begin{aligned}
&= \psi(\gamma_v(1)) \\
&= \psi(q').
\end{aligned}$$

Thus, $\varphi = \psi$ on U . □

3. Local Existence of Statistical Diffeomorphism between Statistical Manifolds

From this section, we shall treat statistical diffeomorphisms. We first state the setting of our theorems.

Setting 3.1. Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two statistical manifolds of the same dimension, and $p \in M$, $\bar{p} \in \bar{M}$. Let also a linear isometry $\Phi : T_p M \rightarrow T_{\bar{p}} \bar{M}$ be given. In addition, choose a convex open subset $\mathcal{U}_p \subset T_p M$ which determines normal coordinate neighborhoods $U := \text{Exp}_p(\mathcal{U}_p)$ and $\bar{U} := \bar{\text{Exp}}_{\bar{p}}(\Phi(\mathcal{U}_p))$ with respect to ∇^g and $\bar{\nabla}^{\bar{g}}$ around p and $\bar{p}$, respectively.

Then we can define a diffeomorphism $\varphi : U \rightarrow \bar{U}$ by

$$\varphi := \bar{\text{Exp}}_{\bar{p}} \circ \Phi \circ (\text{Exp}_p)^{-1}. \quad (3.1)$$

Given a ∇^g -geodesic γ in U emanating from p , $\bar{\gamma} := \varphi \circ \gamma$ is a geodesic relative to $\bar{\nabla}^{\bar{g}}$ in $\bar{U}$ emanating from $\bar{p}$. Denoting by q and $\bar{q} := \varphi(q)$ endpoints of γ and $\bar{\gamma}$, respectively. Then we define a linear isometry $\Phi_\gamma : T_q M \rightarrow T_{\bar{q}} \bar{M}$ by

$$\Phi_\gamma := \bar{P}_{\bar{\gamma}} \circ \Phi \circ (P_\gamma)^{-1}, \quad (3.2)$$

where $P_\gamma : T_p M \rightarrow T_q M$ and $\bar{P}_{\bar{\gamma}} : T_{\bar{p}} \bar{M} \rightarrow T_{\bar{q}} \bar{M}$ are parallel displacements relative to ∇^g and $\bar{\nabla}^{\bar{g}}$ along γ and $\bar{\gamma}$, respectively. It is remarked that they are linear isometries.

Let us consider statistical diffeomorphisms under the setting above. Hereafter, we denote by R^g and $R^{\bar{g}}$ Riemannian curvature tensor fields, and by K and $\bar{K}$ difference tensor fields on (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$, respectively. The following is our main result:

Theorem 3.2. *Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two statistical manifolds of the same dimension, and $p \in M$ and $\bar{p} \in \bar{M}$. Let $\Phi : T_p M \rightarrow T_{\bar{p}} \bar{M}$ be a linear isometry. Set φ and Φ_γ as (3.1) and (3.2) in Setting 3.1. If for any ∇^g -geodesic γ in U starting at p we have*

$$\Phi_\gamma(R^g(u, v)w) = R^{\bar{g}}(\Phi_\gamma(u), \Phi_\gamma(v))\Phi_\gamma(w), \quad (3.3)$$

$$\Phi_\gamma(K(u, v)) = \bar{K}(\Phi_\gamma(u), \Phi_\gamma(v)) \quad (3.4)$$

for any $u, v, w \in T_q M$, then φ is a statistical diffeomorphism from U onto $\bar{U}$.

Conversely, it is easy to see that a statistical diffeomorphism holds (3.3) and (3.4). Thus, we can say that this theorem gives us *necessary and sufficient conditions* for φ defined by (3.1) to be a statistical diffeomorphism.

In order to prove Theorem 3.2, we prepare the following lemma:

Lemma 3.3 [9]. *Let (M, ∇) and $(\bar{M}, \bar{\nabla})$ be two manifolds of the same dimension endowed with torsion-free affine connection, and $p \in M$ and $\bar{p} \in \bar{M}$. Let $\Phi : T_p M \rightarrow T_{\bar{p}} \bar{M}$ be a linear map. Set φ and Φ_γ as (3.1) and (3.2) in Setting 3.1 relative to ∇ and $\bar{\nabla}$. Suppose that for any ∇ -geodesic γ in U starting at p we have*

$$\Phi_\gamma(R(u, v)w) = \bar{R}(\Phi_\gamma(u), \Phi_\gamma(v))\Phi_\gamma(w)$$

for any $u, v, w \in T_q M$, where R and $\bar{R}$ are curvature tensor fields, respectively. Then φ_* coincides with Φ_γ at the endpoint of each γ .

Proof of Theorem 3.2. By using Lemma 3.3, φ_* coincides with Φ_γ at the endpoint of each γ . Thus, we obtain

$$g(X, Y) = \bar{g}(\varphi_*X, \varphi_*Y)$$

for any $X, Y \in \Gamma(TU)$, and then we get $\nabla_{\varphi_*X}^{\bar{g}}\varphi_*Y = \varphi_*(\nabla_X^g Y)$. By assumption (3.4), we have

$$(\varphi_*)_q(K(u, v)) = \bar{K}((\varphi_*)_q(u), (\varphi_*)_q(v))$$

for any $u, v \in T_qM$. We now take $X, Y \in \Gamma(TU)$ such that $X_q = u$, $Y_q = v$, then at $q \in U$,

$$\varphi_*(\nabla_X Y - \nabla_X^g Y) = \bar{\nabla}_{\varphi_*X}\varphi_*Y - \nabla_{\varphi_*X}^{\bar{g}}\varphi_*Y,$$

thus,

$$\bar{\nabla}_{\varphi_*X}\varphi_*Y = \varphi_*(\nabla_X Y).$$

Since $q \in U$ is an arbitrarily point, we conclude $\bar{\nabla}_{\varphi_*X}\varphi_*Y = \varphi_*(\nabla_X Y)$. We have, thus, proved the theorem. $\square$

We notice that $\varphi := \overline{\text{Exp}}_{\bar{p}} \circ \Phi \circ (\text{Exp}_p)^{-1}$ is a unique statistical diffeomorphism from Proposition 2.6. If difference tensor fields K and $\bar{K}$ are vanished, then Theorem 3.2 coincides with a local setting of the Cartan-Ambrose-Hicks theorem (Theorem 1.3).

By using Lemma 2.2, we obtain the following corollary from Theorem 3.2.

Corollary 3.4. *Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two statistical manifolds of the same dimension, and $p \in M$ and $\bar{p} \in \bar{M}$. Let $\Phi : T_pM \rightarrow T_{\bar{p}}\bar{M}$ be a linear isometry. Set φ and Φ_γ as (3.1) and (3.2) in Setting 3.1. If for any ∇^g -geodesic γ in U starting at p we have*

$$\Phi_\gamma(R(u, v)w) = \bar{R}(\Phi_\gamma(u), \Phi_\gamma(v))\Phi_\gamma(w), \quad (3.5)$$

$$\Phi_\gamma(K(u), K(v)) = \bar{K}(\Phi_\gamma(u), \Phi_\gamma(v)), \quad (3.6)$$

$$\Phi_\gamma(\nabla K)(v, w, u) = (\bar{\nabla} \bar{K})(\Phi_\gamma(v), \Phi_\gamma(w); \Phi_\gamma(u)) \quad (3.7)$$

for all $u, v, w \in T_q M$, then ϕ is a statistical diffeomorphism from U onto $\bar{U}$.

Proof. By using (2.1) and assumptions (3.5), (3.6) and (3.7), we obtain

$$\begin{aligned} \Phi_\gamma(R^g(x, y)z) &= \Phi_\gamma(R(x, y)z) - \Phi_\gamma(\nabla K)(y, z; x) + \Phi_\gamma(\nabla K)(x, z; y) \\ &\quad + \Phi_\gamma(K(x, K(y, z))) - \Phi_\gamma(K(y, K(x, z))) \\ &= \bar{R}(\Phi_\gamma(x), \Phi_\gamma(y))\Phi_\gamma(z) \\ &\quad - (\bar{\nabla} \bar{K})(\Phi_\gamma(y), \Phi_\gamma(z); \Phi_\gamma(x)) \\ &\quad + (\bar{\nabla} \bar{K})(\Phi_\gamma(x), \Phi_\gamma(z); \Phi_\gamma(y)) \\ &\quad + K(\Phi_\gamma(x), K(\Phi_\gamma(y), \Phi_\gamma(z))) \\ &\quad - K(\Phi_\gamma(y), K(\Phi_\gamma(x), \Phi_\gamma(z))) \\ &= R^g(\Phi_\gamma(x), \Phi_\gamma(y))\Phi_\gamma(z). \end{aligned}$$

As a direct consequence of this, we obtain (3.3). The given assumption (3.6) coincides with (3.3), thus, completing the proof by virtue of Theorem 3.2. $\square$

In Theorem 3.2, we assume that

$$P_\gamma(R^g(u, v)w) = R^g(P_\gamma(u), P_\gamma(v))P_\gamma(w),$$

$$P_\gamma(K(u, v)) = \bar{K}(P_\gamma(u), P_\gamma(v))$$

for any ∇^g -geodesics γ starting at p . When we consider the behavior of the Riemannian curvature field R^g and the difference tensor field K of a statistical manifold (M, ∇, g) , the simplest case is given by $R^g = 0$ and $K = 0$, where M is locally isometric to a Euclidean space. The next simplest case is $\nabla^g R^g = 0$ and $\nabla^g K = 0$, that is, R^g and K are both parallel relative to ∇^g . These conditions are equivalent to the following:

$$P_c(R^g(u, v)w) = \bar{R}^g(P_c(u), P_c(v))P_c(w),$$

$$P_c(K(u, v)) = \bar{K}(P_c(u), P_c(v))$$

for all piecewise smooth curves c in M . These conditions are extension of assumptions of Theorem 3.2. We remark that $\nabla^g R^g = 0$ is known to be equivalent to the condition that (M, g) is Riemannian locally symmetric space. For example, a Riemannian manifold of constant Riemannian curvature is a Riemannian locally symmetric space. From the discussion above, we consider the following corollary:

Corollary 3.5. *In addition to the assumption of Theorem 3.2, we suppose $\nabla^g R^g = 0$, $\nabla^g K = 0$, $\nabla^{\bar{g}} \bar{R}^g = 0$ and $\nabla^{\bar{g}} \bar{K} = 0$. If Φ holds*

$$\Phi(R^g(u, v)w) = \bar{R}^g(\Phi(u), \Phi(v))\Phi(w), \quad (3.8)$$

$$\Phi(K(u, v)) = \bar{K}(\Phi(u), \Phi(v)) \quad (3.9)$$

for any $u, v, w \in T_p M$, then Φ is a statistical diffeomorphism.

Proof. Since $\nabla^g R^g = 0$, $\nabla^g K = 0$, for any ∇^g -geodesic γ in U emanating from p , we get

$$P_\gamma^g(R^g(u, v)w) = R^g(P_\gamma^g(u), P_\gamma^g(v))P_\gamma^g(w),$$

$$P_\gamma^g(K(u, v)) = \bar{K}(P_\gamma^g(u), P_\gamma^g(v))$$

for any $u, v, w \in T_q M$, where q is the endpoint of γ . Similarly, from $\nabla^{\bar{g}} R^{\bar{g}} = 0$, $\nabla^{\bar{g}} \bar{K} = 0$, we obtain

$$\begin{aligned} P_{\bar{\gamma}}^{\bar{g}}(R^{\bar{g}}(\bar{u}, \bar{v})\bar{w}) &= R^{\bar{g}}(P_{\bar{\gamma}}^{\bar{g}}(\bar{u}), P_{\bar{\gamma}}^{\bar{g}}(\bar{v}))P_{\bar{\gamma}}^{\bar{g}}(\bar{w}), \\ P_{\bar{\gamma}}^{\bar{g}}(K(\bar{u}, \bar{v})) &= \bar{K}(P_{\bar{\gamma}}^{\bar{g}}(\bar{u}), P_{\bar{\gamma}}^{\bar{g}}(\bar{v})) \end{aligned}$$

for any $\bar{u}, \bar{v}, \bar{w} \in T_{\bar{q}} \bar{M}$, where $\bar{\gamma} := \varphi \circ \gamma$ is the $\nabla^{\bar{g}}$ -geodesic from $\bar{p}$ to $\bar{q} := \varphi(q)$. Therefore, from assumptions (3.8) and (3.9), we obtain (3.3) and (3.4). By using Theorem 3.2, we can conclude that φ is a statistical diffeomorphism. $\square$

In general, $\varphi : M \rightarrow \bar{M}$ is said to be a *local diffeomorphism* if for each $p \in M$ there exists a neighborhood U such that $\varphi|_U : U \rightarrow \varphi(U)$ is a diffeomorphism. Furthermore, since φ is a statistical structure preserving, we say that it is a *local statistical diffeomorphism*. When K and $\bar{K}$ are both vanished, φ is a local isometry.

Proposition 3.6. *Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two statistical manifolds, and $\varphi : M \rightarrow \bar{M}$ be a local isometry. For any $p \in M$, if $(\varphi_*)_p$ holds*

$$(\varphi_*)_p(K(u, v)) = \bar{K}((\varphi_*)_p(u), (\varphi_*)_p(v)) \quad (3.10)$$

for any $u, v \in T_p M$, then φ is a local statistical diffeomorphism.

Proof. Since φ is a local isometry, for any $p \in M$, we obtain

$$(\varphi_*)_p(R^g(u, v)w) = R^{\bar{g}}((\varphi_*)_p(u), (\varphi_*)_p(v))(\varphi_*)_p(w)$$

for any $u, v, w \in T_p M$. Take a ∇^g -normal neighborhood $U \subset M$ around p such that $\bar{U} := \varphi(U) \subset \bar{M}$ is a $\nabla^{\bar{g}}$ -normal neighborhood around $\varphi(p)$. In

addition, for any $q \in U$, we take a ∇^g -geodesic γ from p to q . Let P_γ be the ∇^g -parallel translation from p to q along γ , and $\bar{P}_{\varphi \circ \gamma}$ be the $\nabla^{\bar{g}}$ -parallel translation from $\varphi(p)$ to $\varphi(q)$ along $\varphi \circ \gamma$ in $\bar{U}$. Since $(\varphi_*)_q \circ P_\gamma = \bar{P}_{\varphi \circ \gamma} \circ (\varphi_*)_p$, we see that $\Phi_\gamma = (\varphi_*)_q$. Thus, we obtain

$$\Phi_\gamma(R^g(u, v)w) = R^{\bar{g}}(\Phi_\gamma(u), \Phi_\gamma(v))\Phi_\gamma(w)$$

for any $u, v, w \in T_q M$. In addition, by using assumption (3.10), we have

$$\Phi_\gamma(K(u, v)) = \bar{K}(\Phi_\gamma(u), \Phi_\gamma(v)).$$

By virtue of Theorem 3.2, we can conclude that φ is a local statistical diffeomorphism. $\square$

4. Local Existence of Statistical Diffeomorphism between Hessian Manifolds

In the preceding section, we gave an idea to the local existence of statistical diffeomorphism between statistical manifolds, and showed Theorem 3.2. In this section, we shall treat Hessian manifolds, which are the important subclass of statistical manifolds, and consider the local existence of statistical diffeomorphism between Hessian manifolds under Setting 3.1. We denote by H and $\bar{H}$ the Hessian curvature tensor fields on M and $\bar{M}$, respectively.

Theorem 4.1. *Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two Hessian manifolds of the same dimension, and $p \in M$ and $\bar{p} \in \bar{M}$. Let $\Phi : T_p M \rightarrow T_{\bar{p}} \bar{M}$ be a linear isometry. Set φ and Φ_γ as (3.1) and (3.2) in Setting 3.1. If for any ∇^g -geodesic γ in U starting at p we have*

$$\Phi_\gamma(K(u, v)) = \bar{K}(\Phi_\gamma(u), \Phi_\gamma(v)),$$

$$\Phi_\gamma(H(u, v)w) = \bar{H}(\Phi_\gamma(u), \Phi_\gamma(v))\Phi_\gamma(w)$$

for all $u, v, w \in T_q M$, then φ is a statistical diffeomorphism from U onto $\bar{U}$.

Proof. By Definition 1.4, we see that R and $\bar{R}$ are vanished. Since $H(X, Y)Z = -(\nabla K)(Y, Z; X)$, we obtain Theorem 4.1 from Corollary 3.4.

□

Next, we consider the case where Hessian manifolds are of constant Hessian sectional curvature.

Corollary 4.2. *Let (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ be two Hessian manifolds of the same dimension and of constant Hessian sectional curvature $c \in \mathbb{R}$. Now, let $p \in M$ and $\bar{p} \in \bar{M}$ be two any fixed points, and φ an isometry between their neighborhoods U and $\bar{U}$ such that $\varphi(p) = \bar{p}$. If for any $q \in U$ we have*

$$(\varphi_*)_q(K(u, v)) = \bar{K}((\varphi_*)_q(u), (\varphi_*)_q(v))$$

for all $u, v \in T_q M$, then φ is a statistical diffeomorphism.

Proof. From Proposition 2.4, both (M, ∇, g) and $(\bar{M}, \bar{\nabla}, \bar{g})$ are of constant Riemannian curvature $-c/4 \in \mathbb{R}$. Now, it is to be noted that a Riemannian manifold of constant Riemannian curvature is a Riemannian locally symmetric space, that is, $\nabla^g R^g = 0$. Hence, by using Theorem 1.3, we see that for any points $p \in M$ and $\bar{p} \in \bar{M}$, there exists an isometry φ between their neighborhoods. By virtue of Proposition 3.6, we arrive at the conclusion.

□

Finally, let us consider some examples of statistical diffeomorphisms.

Example 4.3. We define $((\mathbb{R}^+)^n, \nabla, g)$ by

$$(\mathbb{R}^+)^n := \{y = {}^t(y^1, \dots, y^n) \in \mathbb{R}^n \mid y^1 > 0, \dots, y^n > 0\},$$

$$\nabla_{\frac{\partial}{\partial y^i}} \frac{\partial}{\partial y^j} := -\frac{1}{y^j} \delta_{ij} \frac{\partial}{\partial y^j},$$

$$g := g^E|_{(\mathbb{R}^+)^n} = \sum_{i=1}^n (dy^i)^2,$$

where $g^E = \sum_{i,j=1}^n \delta_{ij} dy^i dy^j$. Then $((\mathbb{R}^+)^n, \nabla, g)$ is a Hessian manifold of constant Hessian sectional curvature 0. The triplet

$$\left(\mathbb{R}^n, D, \tilde{g} = Dd \left(\sum_{i=1}^n e^{2x^i} / 4 \right) \right)$$

is also a Hessian manifold of constant Hessian sectional curvature 0, where D is the standard connection. We define $\varphi : (\mathbb{R}^+)^n \rightarrow \mathbb{R}^n$ by

$$\varphi(y^1, \dots, y^n) := (\log y^1, \dots, \log y^n),$$

then φ is a statistical diffeomorphism.

Example 4.4. Let $(\mathbb{H}^n, g^{\mathbb{H}})$ be the upper half space of constant curvature -1 defined by

$$\mathbb{H}^n := \{y = (y^1, \dots, y^n) \in \mathbb{R}^n \mid y^n > 0\},$$

$$g^{\mathbb{H}} := \frac{\sum_{i=1}^n dy^i dy^i}{(y^n)^2}.$$

Now, we define a flat affine connection on $\mathbb{H}^n$ as follows:

$$\nabla_{\frac{\partial}{\partial y^n}} \frac{\partial}{\partial y^n} := \frac{1}{y^n} \frac{\partial}{\partial y^n},$$

$$\nabla_{\frac{\partial}{\partial y^i}} \frac{\partial}{\partial y^j} := \frac{2}{y^n} \delta_{ij} \frac{\partial}{\partial y^n},$$

$$\nabla^{\mathbb{H}} \frac{\partial}{\frac{\partial}{\partial y^i} \partial y^n} := \nabla^{\mathbb{H}} \frac{\partial}{\frac{\partial}{\partial y^n} \partial y^j} := 0,$$

where $i, j = 1, \dots, n-1$. Then $(\mathbb{H}^n, \nabla^{\mathbb{H}}, g^{\mathbb{H}})$ is a Hessian manifold of constant Hessian sectional curvature 4. It should be remarked that $(\mathbb{H}^2, \nabla^{\mathbb{H}}, g^{\mathbb{H}})$ expresses the statistical model of one-dimensional normal distributions.

Let $\varphi : \mathbb{H}^n \rightarrow \mathbb{H}^n$ be a diffeomorphism. Then φ is a statistical diffeomorphism if and only if

$$\varphi(y) = \left(c \sum_{j=1}^{n-1} a_j^1 y^j + b^1, \dots, c \sum_{j=1}^{n-1} a_j^{n-1} y^j + b^{n-1}, cy^n \right), \quad (4.1)$$

where $A = [a_j^i]_{i,j=1,\dots,n-1} \in O(n-1)$, $b \in \mathbb{R}^{n-1}$, $c > 0$. In particular, when $n = 2$, we have

$$\varphi(y^1, y^2) = (\pm cy^1 + b, cy^2), \quad (4.2)$$

where $b, c \in \mathbb{R}$, $c > 0$. Therefore, we see that a statistical diffeomorphism on statistical models of one-dimensional normal distributions is given in (4.2).

This can be seen in the following way. Let φ be a statistical diffeomorphism. Then the following hold:

$$\nabla^{\mathbb{H}} \frac{\partial}{\varphi_* \frac{\partial}{\partial y^s}} \varphi_* \frac{\partial}{\partial y^t} - \varphi_* \left(\nabla^{\mathbb{H}} \frac{\partial}{\partial y^s} \frac{\partial}{\partial y^t} \right) = 0, \quad (4.3)$$

$$\nabla^g \frac{\partial}{\varphi_* \frac{\partial}{\partial y^s}} \varphi_* \frac{\partial}{\partial y^t} - \varphi_* \left(\nabla^g \frac{\partial}{\partial y^s} \frac{\partial}{\partial y^t} \right) = 0, \quad (4.4)$$

where $s, t = 1, \dots, n$. We note that the Levi-Civita connection of $g^{\mathbb{H}}$ is given by

$$\nabla_{\frac{\partial}{\partial y^n}}^g \frac{\partial}{\partial y^n} = -\frac{1}{y^n} \frac{\partial}{\partial y^n},$$

$$\nabla_{\frac{\partial}{\partial y^i}}^g \frac{\partial}{\partial y^j} = \frac{1}{y^n} \delta_{ij} \frac{\partial}{\partial y^n},$$

$$\nabla_{\frac{\partial}{\partial y^i}}^g \frac{\partial}{\partial y^n} = \nabla_{\frac{\partial}{\partial y^n}}^g \frac{\partial}{\partial y^j} = -\frac{1}{y^n} \frac{\partial}{\partial y^i},$$

where $i, j = 1, \dots, n-1$. Now, we write $\varphi(y) = (\varphi^1(y), \dots, \varphi^n(y))$ for $y \in \mathbb{H}^n$. From (4.3) and (4.4) with direct calculation, we obtain

$$\frac{\partial^2 \varphi^k}{\partial y^i \partial y^j}(y) = 0, \quad (4.5)$$

$$\frac{\partial^2 \varphi^n}{\partial y^n \partial y^n}(y) = 0, \quad (4.6)$$

$$\frac{\partial \varphi^n}{\partial y^n}(y) = \frac{1}{y^n} \varphi^n(y), \quad (4.7)$$

$$\sum_{l=1}^{n-1} \frac{\partial \varphi^l}{\partial y^i}(y) \frac{\partial \varphi^l}{\partial y^j}(y) = \frac{1}{(y^n)^2} (\varphi^n)^2 \delta_{ij}, \quad (4.8)$$

where $i, j, k = 1, \dots, n-1$. Indeed, (4.5) and (4.6) are obtained from (4.3) or (4.4). On the other hand, (4.7) and (4.8) are yielded by combining (4.3) with (4.4). From (4.5), we see that $\varphi^k(y)$ is a linear function, that is, $\varphi^k(y) = \sum_{i=1}^{n-1} d_i^k y^i + b^k$, where $d_i^k, b^k \in \mathbb{R}$ ($k = 1, \dots, n-1$). By using both (4.6) and (4.7), we get $\varphi^n(y) = cy^n$, where $c > 0$. Moreover, from

(4.8), we obtain

$$\sum_{l=1}^{n-1} a_i^l a_j^l = \delta_{ij}, \quad (4.9)$$

where $a_i^l := d_i^l / c$. Expression (4.9) can be restated that $A = [a_j^i]_{i,j=1,\dots,n-1} \in O(n-1)$. For the above reasons, we reach (4.1). Conversely, we set ϕ as (4.1), and see that ϕ is a statistical diffeomorphism by computation.

It is known that the routes of ∇^g -geodesics in $\mathbb{H}^2$ are the constant speed parametrizations of all vertical lines y^1 constant (for a while, we name them L-type ∇^g -geodesics) and all semicircles with centers on the y^1 -axis (C-type ∇^g -geodesics), and L-type ∇^g -geodesics and C-type ∇^g -geodesics can translate each other by isometric transformations. However, in the case of statistical diffeomorphisms, L-type ∇^g -geodesics and C-type ∇^g -geodesics cannot translate each other.

Acknowledgments

The author would like to express his sincere thanks to Professor Furuhata for his useful advice.

This work was supported by JSPS KAKENHI Grant Number JP18K18714.

References

- [1] S. Amari, Information Geometry and its Applications, Appl. Math. Sci., Springer, Vol. 194, 2016.
- [2] S. Amari and H. Nagaoka, Method of Information Geometry, Amer. Math. Soc., Providence, Oxford University Press, Oxford, 2000.

- [3] W. Ambrose, Parallel translation of Riemannian curvature, *Ann. Math.* 64 (1956), 337-363.
- [4] N. Ay, J. Jost, H. V. Le and L. Schwachhöfer, *Information Geometry*, Springer, 2017.
- [5] E. Cartan, *Géométrie des espaces de Riemann*, Gauthier-Villars, Paris, 1946.
- [6] H. Furuhashi, Hypersurfaces in statistical manifolds, *Differ. Geom. Appl.* 27 (2009), 420-429.
- [7] H. Furuhashi, Geometry of statistical submanifolds, slides of the presentation at PADGE2012, KU Leuven.
<https://wis.kuleuven.be/events/archive/padge2012/slides/furuhashi.pdf>.
- [8] H. Furuhashi and T. Kurose, Hessian manifolds of nonpositive constant Hessian sectional curvature, *Tohoku Math. J.* 65 (2013), 31-42.
- [9] N. J. Hicks, Connexion preserving spray maps, *Illinois J. Math.* 10 (1966), 661-679.
- [10] T. Kurose, Dual connections and affine geometry, *Math. Z.* 203 (1990), 115-121.
- [11] T. Kurose, On the divergences of 1-conformally flat statistical manifolds, *Tohoku Math. J.* 46 (1994), 427-433.
- [12] S. L. Lauritzen, *Statistical manifolds, Differential Geometry in Statistical Inferences*, IMS Lecture Notes Monograph Series 10, Institute of Mathematical Statistics, Hayward California, 1987, pp. 96-163.
- [13] K. Nomizu and T. Sasaki, *Affine Differential Geometry*, Cambridge University Press, 1994.
- [14] B. O'Neill, *Semi-Riemannian Geometry with Applications to Relativity*, Academic Press, 1983.
- [15] H. Shima, *The Geometry of Hessian Structure*, World Scientific Publ., 2007.
- [16] M. Takeuchi, *Lie Groups II*, Transl. Math. Monogr., Amer. Math. Soc., Vol. 85, 1991, pp. 113-260.
- [17] J. A. Wolf, *Spaces of Constant Curvature*, 6th ed., Amer. Math. Soc., 2011.

Naoto Satoh: n-satoh@math.sci.hokudai.ac.jp