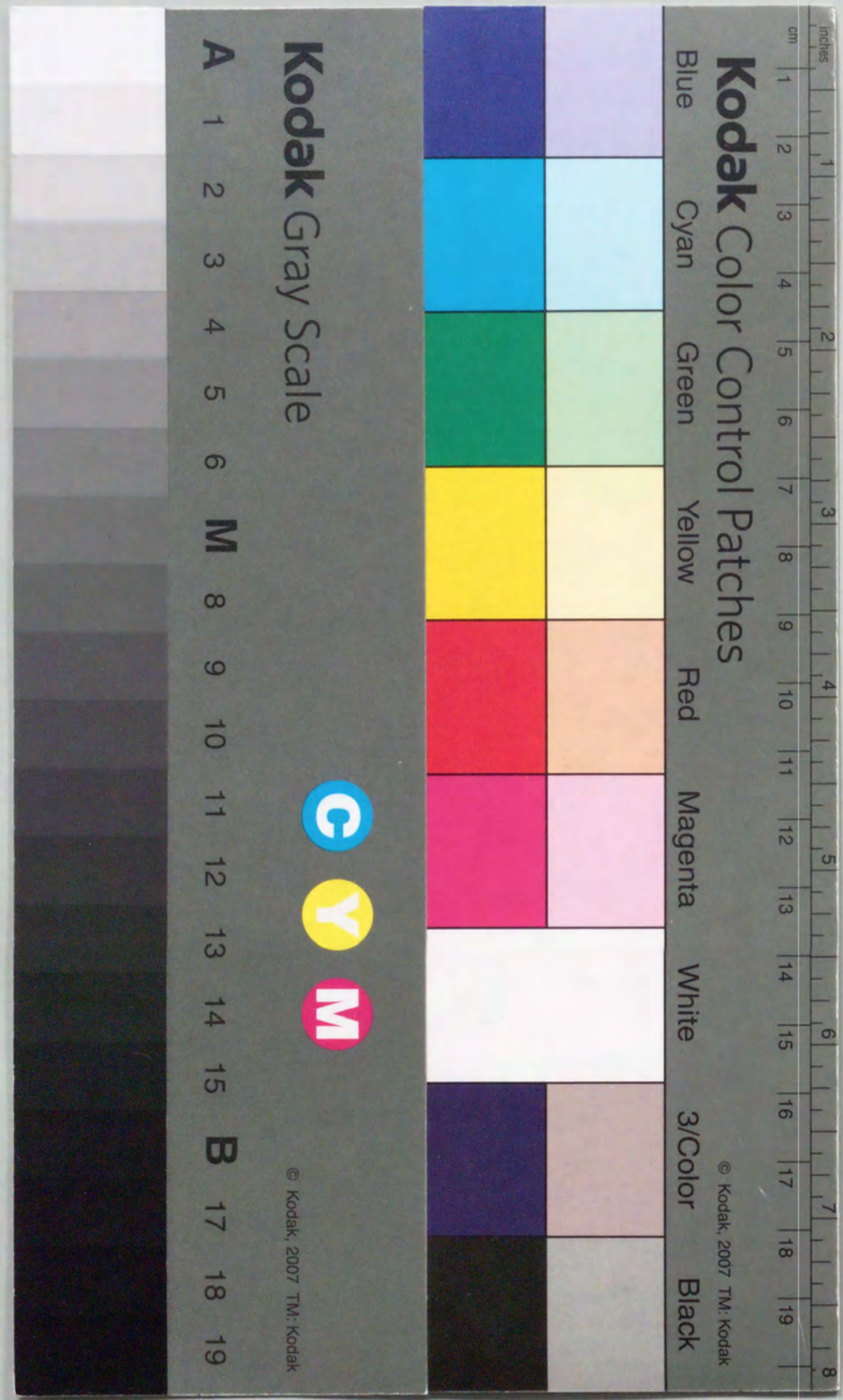

Conformal Flatness and Autoparallelism of Statistical Submanifolds

Keiko Uohashi

Graduate School of Engineering Science
Osaka University

1999

CONFORMAL FLATNESS AND AUTOPARALLELISM OF STATISTICAL SUBMANIFOLDS

A Dissertation Submitted to
the Graduate School of Engineering Science
of Osaka University

by
Keiko UOHASHI

January 1999

Preface

Statistical manifolds have been studied in terms of information geometry, which is investigation of manifolds with dualistic structures and its applications. Dualistic structures appear in not only statistical inference but also many other engineering applications. Consideration of statistical submanifolds is very convenient for above studies. In this paper, we investigate particular statistical submanifolds: conformally flat statistical submanifolds and doubly autoparallel statistical submanifolds.

A convex function on an affine domain generates the Riemannian metric. The domain with the metric is called a Hessian domain. The function also generates the dual affine connection and decides the dually flat structure of the domain. Thus the Hessian domain is a flat statistical manifold. For a Riemannian metric h and a function ϕ, we call $e^{\phi}h$ a conformal transformation of h. Moreover considering deformation of affine connections, Okamoto, Amari and Takeuchi first treated the concept of α-conformal equivalence with respect to sequential estimation theory. An α-conformally equivalent statistical manifold to a flat statistical manifold is said to be α-conformally flat. As $\alpha = 1$, 1-conformally flat statistical manifolds can be realized in the affine space of codimension one. However existence of 1-conformally flat statistical *sub*manifold was not known. Then we shall prove that a level surface of the function is a 1-conformally flat statistical submanifold.

It is known that there exist the canonical divergences of a flat statistical manifold and of a 1-conformally flat statistical manifold. Divergences are pseudo-distance treated first in information theory and statistics. In this paper we consider a foliation defined by level surfaces and its orthogonal foliation, and investigate divergences restricted to leaves of these foliations. Next we give the decomposition of the divergence of a Hessian domain with respect to orthogonal foliations. On a flat statistical manifold, Nagaoka and Amari gave a decomposition of the divergence known as the extended Pythagorean theorem, considering autoparallel submanifolds. We obtain a decomposition different from their one, and show the application to gradient systems.

Finally we study dually flat structures on symmetric cones associated with Jordan algebras. A symmetric cone has the characteristic function, which generates the canonical dually flat structure on it. We give an interpretation of connections, a geometrical concept, in terms of Jordan algebras and show relation with doubly autoparallel submanifolds and Jordan subalgebras. We can solve semidefinite programming (SDP), convex programming on a cone of positive definite symmetric matrices, without iterations of Newton-method, if its feasible region is doubly autoparallel submanifold of the cone. Then it is useful for convex programming on symmetric cones to know conditions for doubly autoparallelism of submanifolds.

Acknowledgement

I would like to express sincere thanks to Professor Takao Fujii for his guidance on control systems theory. I am thankful to Dr. Atsumi Ohara for his advice on information geometric approach to control systems theory and convex programming. I would like to thank Professor Hirohiko Shima in Yamaguchi University, who gave me valuable comments on Jordan algebras and Hessian manifolds related information geometry. I also would like to thank Dr. Takashi Kurose in Fukuoka University, who discussed with me dualistic structures of statistical manifolds.

I am thankful to Professor Satoru Fujishige and Professor Yoshimasa Nakamura, who are members of my dissertation committee. They read this dissertation and gave me valuable suggestions before its completion.

Professor Nobuhide Suda in Hosei University invited me to study mathematical control systems theory. I would like to give him my special thanks. This dissertation could not be completed without him. I am grateful very much to Dr. Takashi Shimomura and Dr. Taro Tsujino, who advised and encouraged me in seminars.

Finally I would like to express appreciation to Ms. Tokiko Nakagawa, Ms. Kyoko Nakamura and all the members of Fujii laboratory at Graduate School of Engineering Science of Osaka University.

Contents

Chapter 1

Introduction

1.1 Backgrounds and motivations

Statistical manifolds have been studied in terms of information geometry. Dualistic structures of statistical manifolds play important roles on statistical inference, control systems theory, convex programming, and so on ([A][OA][OSA]). Especially a dually flat structure is a Hessian structure, whose metric is led by the second order differentiation of a convex function. Information geometry and theory of Hessian structure were studied first independently, but now relation with them are pointed out: for example it becomes known that a Hessian structure gives geometry of an exponential family ([SH]). However they have not been studied deeply. Thus we shall investigate statistical manifolds using theory of Hessian structures.

Applications of the dually flat structures of submanifolds are in [FA][OSA]. Embedding curvatures of non-flat statistical submanifolds are related with various estimators in statistics. From a differential geometrical point of view, non-flat statistical manifolds are studied in [K1][K2][Mz], but we see few results on them as statistical submanifolds without dually flat structures. Then we treat non-flat dualistic structures on submanifolds, especially on level surfaces of Hessian domains, and obtain properties of flat statistical manifolds.

Symmetric cones are typical Hessian domains, since characteristic functions of symmetric cones generate the canonical dually flat structure on it. We describe dual affine connections on symmetric cones in terms of the Jordan algebra. Shima gave correspondence between Jordan algebras and affine connections on symmetric spaces with Hessian structure ([S1][S3]). Ohara presented a relation between a Jordan algebra and the dual connection especially on a cone of positive definite symmetric matrices ([O]). [S1][S3] make use of vector fields induced by Lie algebras, and [O] treats of vector fields along the canonical affine coordinates. In this paper, we shall generalize the result in [O] to other symmetric cones.

Convex programming, e.g., semidefinite programming (SDP) is useful for system and control theory, and so on ([F][O]). We can solve SDP without iterations of Newton-method, if its feasible region is doubly autoparallel submanifold of the cone. Then we see doubly autoparallel submanifolds of symmetric cones.

1.2 Overview

1.2.1 Statistical manifolds

First we give foundations and notations of statistical manifolds. In Chapter 2, there are definitions of statistical manifolds, statistical submanifolds, α-conformal flatness, Hessian domains, dual statistical manifolds, and Legendre transform.

1.2.2 Conformal flatness of level surfaces

We study 1-conformal flatness and α-conformal equivalence of level surfaces in flat statistical manifolds in Chapter 3. Let φ be a convex function on a domain Ω in an affine space $\mathbf{A}^{n+1}$. Denoting by $\tilde{D}$ the canonical flat affine connection on $\mathbf{A}^{n+1}$, we can consider a Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ a flat statistical manifold.

In Section 3.1, we show that, if $\tilde{g}$ is positive definite, n-dimensional level surfaces of φ are 1-conformally flat statistical submanifolds of $(\Omega, \tilde{D}, \tilde{g})$, and that a 1-conformally flat statistical manifold with a Riemannian metric can be locally realized as a statistical submanifold of a flat statistical manifold.

In Section 3.2, we give a procedure to realize a statistical manifold, which is α-conformally equivalent to a manifold with an α-connection, as a statistical submanifold.

1.2.3 Divergences

We construct dual foliations: $\mathcal{F}$ defined by n-dimensional level surfaces and $\mathcal{F}^{\perp}$ orthogonal to $\mathcal{F}$, and discuss divergences on leaves of the foliations.

Projective transform is dilation of surfaces in an affine space. In Subsection 4.1.1, we see that the dual connection D' of D has projective change into a flat connection, that is, D is dual-projectively flat, where D is the induced affine connection on a level surface. Next, we show that dual-projectively

equivalent affine connections can be led on a leaf of a foliation $\mathcal{F}$ in Subsection4.2.1. In addition we show that a leaf of $\mathcal{F}^{\perp}$ is a $\tilde{D}'$-geodesic, where $\tilde{D}'$ is the dual connection of $\tilde{D}$.

We also discuss divergences on leaves of the foliations $\mathcal{F}$ and $\mathcal{F}^{\perp}$ in Section 4.2. Nagaoka and Amari first studied divergences of flat statistical manifolds in view of statistics ([A][AN]). Kurose defined the canonical divergences of 1-conformally flat statistical manifolds ([K2]). We show that, for $M \in \mathcal{F}$, Kurose's divergence of a 1-conformally flat statistical submanifold (M, D, g) is the restriction of Nagaoka and Amari's divergence of $(\Omega, \tilde{D}, \tilde{g})$ in Subsection 4.2.1. We give the decomposition of the divergence of $(\Omega, \tilde{D}, \tilde{g})$ with respect to orthogonal foliations $\mathcal{F}$ and $\mathcal{F}^{\perp}$ in Subsection 4.2.2. Next we see that the projection of a point in Ω to M along a leaf of $\mathcal{F}^{\perp}$ is given by minimization of the divergence. Last in Chapter 4, we give a gradient system using the divergence. Gradient systems are important to study relation with information geometry and integrable dynamical systems ([FA][N]).

1.2.4 Autoparallel submanifolds of symmetric cones

In Chapter 5, we describe dualistic structures of symmetric cones and of autoparallel submanifolds. We give notation and formulas on Jordan algebras in Subsection 5.1.1. Next, we introduce dually flat structures on symmetric cones associated with simple Euclidean Jordan algebras in Subsection 5.1.2. We derive a relation between the dual connections and mutations of Jordan algebras in Subsection 5.1.3. For some condition, we prove that submanifolds of symmetric cones are doubly autoparallel if and only if their tangent spaces are Jordan subalgebras, in Subsection 5.2.1. Finally for the same condition, we show that submanifolds are autoparallel with respect to α-connections also if and only if their tangent spaces are Jordan subalgebras, in Subsection 5.2.2.

In the last chapter, we note conclusions.

Chapter 2

Preliminary

2.1 Statistical manifolds

For a torsion-free affine connection ∇ and a pseudo-Riemannian metric h on a manifold N, the triple (N, ∇, h) is called a statistical manifold if ∇h is symmetric. If the curvature tensor R of ∇ vanishes, (N, ∇, h) is said to be flat.

For a statistical manifold (N, ∇, h), let ∇' be an affine connection on N such that

$$Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla'_X Z) \quad \text{for } X, Y, Z \in \mathcal{X}(N),$$

where $\mathcal{X}(N)$ is the set of all tangent vector fields on N. The affine connection ∇' is torsion free, and $\nabla' h$ symmetric. Then ∇' is called the dual connection of ∇, the triple (N, ∇', h) the dual statistical manifold of (N, ∇, h), and (∇, ∇', h) the dualistic structure on N. The curvature tensor of ∇' vanishes if and only if one of ∇ does, and then (∇, ∇', h) is called the dually flat structure.

2.2 Statistical submanifolds

For a pseudo-Riemannian manifold $(\tilde{N}, \tilde{h})$ and a submanifold N of $\tilde{N}$, we call (N, ∇, h) a statistical submanifold of $(\tilde{N}, \tilde{h})$ if (N, ∇, h) is a statistical manifold, where ∇ is an affine connection on N and h the induced pseudo-Riemannian metric for $\tilde{h}$. Let $\tilde{\nabla}$ be an affine connection on $\tilde{N}$. We denote by $T_pN \oplus T_pN^{\perp}$ the orthogonal decomposition of $T_p\tilde{N}$ with respect to $\tilde{h}$, where $T_p\tilde{N}$ and T_pN are the set of all tangent vectors at x on $\tilde{N}$ and on N, respectively. If $(\nabla_X Y)_p$ is the T_pN-component of $(\tilde{\nabla}_X Y)_p$ for $X, Y \in \mathcal{X}(N)$

and an arbitrary x in N, we call (N, ∇, h) the statistical submanifold realized in $(\tilde{N}, \tilde{\nabla}, \tilde{h})$.

If $(\tilde{N}, \tilde{\nabla}, \tilde{h})$ is a statistical manifold for a Riemannian metric $\tilde{h}$ and a submanifold N of $\tilde{N}$, (N, ∇, h) is a statistical manifold for the above induced connection ∇ and the induced metric h ([A][Vo]). For a pseudo-Riemannian metric $\tilde{h}$, (N, ∇, h) is a statistical manifold if h is non-degenerate. Then we often call a statistical submanifold realized in a statistical manifold $(\tilde{N}, \tilde{\nabla}, \tilde{h})$, simply, a statistical submanifold of $(\tilde{N}, \tilde{\nabla}, \tilde{h})$.

Let (N, ∇', h) be the dual statistical manifold of (N, ∇, h). If (N^s, ∇^s, h^s) and $(N^s, \nabla^{s\prime}, h^s)$ are statistical submanifolds of (N, ∇, h) and (N, ∇', h), respectively, $(N^s, \nabla^{s\prime}, h^s)$ is the dual statistical manifold of (N^s, ∇^s, h^s).

2.3 α-conformal flatness

For $\alpha \in \mathbf{R}$, statistical manifolds (N, ∇, h) and $(N, \bar{\nabla}, \bar{h})$ are said to be α-conformally equivalent if there exists a function ϕ on N such that

$$\begin{aligned} \bar{h}(X,Y) &= e^{\phi} h(X,Y), \\ h(\bar{\nabla}_X Y, Z) &= h(\nabla_X Y, Z) - \frac{1+\alpha}{2} d\phi(Z) h(X,Y) \\ &\quad + \frac{1-\alpha}{2}\{d\phi(X) h(Y,Z) + d\phi(Y) h(X,Z)\} \end{aligned}$$

for $X, Y, Z \in \mathcal{X}(N)$. A statistical manifold (N, ∇, h) is called α-conformally flat if (N, ∇, h) is locally α-conformally equivalent to a flat statistical manifold. Statistical manifolds (N, ∇, h) and $(N, \bar{\nabla}, \bar{h})$ are α-conformally equivalent if and only if the dual statistical manifolds (N, ∇', h) and $(N, \bar{\nabla}', \bar{h})$ are $(-\alpha)$-conformally equivalent. Especially, a statistical manifold (N, ∇, h) is 1-conformally flat if and only if the dual statistical manifold (N, ∇', h) is (-1)-conformally flat ([K2]).

2.4 Hessian domains

Let $\tilde{D}$ and $\{x^1, \ldots, x^{n+1}\}$ be the canonical flat affine connection and the canonical affine coordinate system on $\mathbf{A}^{n+1}$, i.e., $\tilde{D}dx^i = 0$. If the Hessian $\tilde{D}d\varphi = \sum_{i,j} \frac{\partial^2 \varphi}{\partial x^i \partial x^j} dx^i dx^j$ is non-degenerate for a function φ on a domain Ω in $\mathbf{A}^{n+1}$, we call $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ a Hessian domain.

A Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ is a flat statistical manifold. Conversely, a flat statistical manifold is locally a Hessian domain ([A][S2]).

2.5 Legendre transform

Let $\mathbf{A}^*_{n+1}$ and $\{x^*_1, \ldots, x^*_{n+1}\}$ be the dual affine space of $\mathbf{A}^{n+1}$ and the dual affine coordinate system of $\{x^1, \ldots, x^{n+1}\}$, respectively. We define the gradient mapping $\tilde{\iota}$ from Ω to $\mathbf{A}^*_{n+1}$ by

$$x^*_i \circ \tilde{\iota} = -\frac{\partial \varphi}{\partial x^i}, \tag{2.1}$$

and a flat affine connection $\tilde{D}'$ on Ω by

$$\tilde{\iota}_*(\tilde{D}'_{\tilde{X}} \tilde{Y}) = \tilde{D}^*_{\tilde{X}} \tilde{\iota}_*(\tilde{Y}) \quad \text{for} \quad \tilde{X}, \tilde{Y} \in \mathcal{X}(\Omega),$$

where $\tilde{D}^*_{\tilde{X}} \tilde{\iota}_*(\tilde{Y})$ is covariant derivative along $\tilde{\iota}$ induced by the canonical flat affine connection $\tilde{D}^*$ on $\mathbf{A}^*_{n+1}$. Then $(\Omega, \tilde{D}', \tilde{g})$ is the dual statistical manifold of $(\Omega, \tilde{D}, \tilde{g})$. We set $x'_i = x^*_i \circ \tilde{\iota} = -\frac{\partial \varphi}{\partial x^i}$. Then $\{x'_1, \ldots, x'_{n+1}\}$ is the affine flat coordinate system with respect to $\tilde{D}'$, i.e., $\tilde{D}'dx'_i = 0$. Remark that a straight line with respect to an affine coordinate $\{x^1, \ldots, x^{n+1}\}$ (resp. $\{x'_1, \ldots, x'_{n+1}\}$) is a $\tilde{D}$- (resp. $\tilde{D}'$-) geodesic, where we call a geodesic relative to $\tilde{D}$ (resp. $\tilde{D}'$) a $\tilde{D}$- (resp. $\tilde{D}'$-) geodesic.

If $\tilde{\iota}$ is invertible, we can define a function on $\Omega^* = \tilde{\iota}(\Omega)$ called the Legendre transform φ^* of φ by

$$\varphi^* \circ \tilde{\iota} = -\sum_i x^i x'_i - \varphi.$$

The triple $(\Omega^*, \tilde{D}', \tilde{g}' = \tilde{D}'d\varphi^*)$ is a flat statistical manifold.

Chapter 3

Dualistic structures on level surfaces

3.1 1-conformal flatness of level surfaces

We give two theorems of level surfaces in Subsection 3.1.1 and prove them in Subsection 3.1.2 and 3.1.3.

3.1.1 Theorems

We obtain the next theorems.

Theorem 3.1. *Let M be a simply connected n-dimensional level surface of φ on an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ with a Riemannian metric $\tilde{g}$, and suppose that $n \geq 2$. If we consider $(\Omega, \tilde{D}, \tilde{g})$ a flat statistical manifold, (M, D, g) is a 1-conformally flat statistical submanifold of $(\Omega, \tilde{D}, \tilde{g})$, where we denote by D and g the connection and the Riemannian metric on M induced by $\tilde{D}$ and $\tilde{g}$.*

Theorem 3.2. *An arbitrary 1-conformally flat statistical manifold of* $\dim n \geq 2$ *with a Riemannian metric can be locally realized as a submanifold of a flat statistical manifold of* $\dim(n+1)$.

3.1.2 Statistical manifolds and affine differential geometry

For a real number c, we call a subset $M := \{p \in \Omega \mid \varphi(p) = c\}$ is called a level surface of φ on Ω. We study a level surface M of φ on an $(n+1)$-dimensional

Hessian domain $(\Omega, \tilde{D}, \tilde{g})$, using affine differential geometry and the concept of statistical submanifolds. A level surface M of φ is an n-dimensional submanifold of Ω if and only if $d\varphi_p \neq 0$ for all $p \in M$. Henceforward, we suppose that $n \geq 2$, that $\tilde{g}$ is a Riemannian metric, and that $d\varphi_p \neq 0$ for all $p \in M$.

Let $\tilde{E}$ be the gradient vector field on Ω defined by

$$\tilde{g}(\tilde{X}, \tilde{E}) = d\varphi(\tilde{X}) \quad \text{for } \tilde{X} \in \mathcal{X}(\Omega).$$

Since $\tilde{g}$ is positive definite and $d\varphi_p \neq 0$ for all $p \in M$, $d\varphi(\tilde{E})$ does not vanish on M and a vector $\tilde{E}_p$ is vertical to T_pM with respect to $\tilde{g}$, where T_pM is the set of all tangent vectors at p on M. We set

$$E = -d\varphi(\tilde{E})^{-1}\tilde{E} \tag{3.1}$$

on M. Then the vector field $\tilde{E}$ is transversal to M, and so is E.

Let x be a canonical immersion of M into Ω. For $\tilde{D}$ and an affine immersion (x, E), we can define the induced affine connection D^E, the fundamental form g^E, the shape operator S^E and the transversal connection form τ^E on M by

$$\tilde{D}_X Y = D^E_X Y + g^E(X, Y)E \tag{3.2}$$

$$\tilde{D}_X E = S^E(X) + \tau^E(X)E \quad \text{for } X, Y \in \mathcal{X}(M). \tag{3.3}$$

We denote by (M, D, g) the statistical submanifold of $(\Omega, \tilde{D}, \tilde{g})$, considering $(\Omega, \tilde{D}, \tilde{g})$ a statistical manifold. Then the next holds.

Lemma 3.3. *A statistical submanifold (M, D, g) coincides with a manifold (M, D^E, g^E) induced by an affine immersion (x, E), i.e.,*

$$D = D^E, \quad g = g^E \quad \text{on } M.$$

Proof. Let $D^{\tilde{E}}$ be the induced affine connection, $g^{\tilde{E}}$ the fundamental form, $S^{\tilde{E}}$ the shape operator, and $\tau^{\tilde{E}}$ the transversal connection form, for $\tilde{D}$ and $\tilde{E}$. Since E_p and $\tilde{E}_p$ are vertical to T_pM for $p \in M$ with respect to $\tilde{g}$, $D = D^E = D^{\tilde{E}}$ holds. From (3.2) and $\tilde{D}_X Y = D^{\tilde{E}}_X Y + g^{\tilde{E}}(X, Y)\tilde{E}$, we have

$$g^{\tilde{E}} = -d\varphi(\tilde{E})^{-1} g^E. \tag{3.4}$$

By [HS] we know that

$$g^{\tilde{E}} = -d\varphi(\tilde{E})^{-1} g. \tag{3.5}$$

From (3.4) and (3.5) $g = g^E$ holds. □

Since g is non-degenerate, so is g^E. Then (x, E) is called a non-degenerate immersion. Moreover, the immersion (x, E) has the following property.

Lemma 3.4. *An affine immersion (x, E) is equiaffine, i.e.,*

$$\tau^E = 0 \quad \text{on } M.$$

Proof. We have

$$\tau^{\tilde{E}} = (d \log \mid d\varphi(\tilde{E}) \mid)(X) \tag{3.6}$$

by [HS]. Calculating the right-hand side of (3.6), we have

$$\tau^{\tilde{E}} = d\varphi(\tilde{E})^{-1} X(d\varphi(\tilde{E})).$$

Thus, we obtain

$$\begin{aligned} \tilde{D}_X E &= -\tilde{D}_X(d\varphi(\tilde{E})^{-1}\tilde{E}) \\ &= -X(d\varphi(\tilde{E})^{-1})\tilde{E} - d\varphi(\tilde{E})^{-1} D_X \tilde{E} \\ &= d\varphi(\tilde{E})^{-2} X(d\varphi(\tilde{E}))\tilde{E} - d\varphi(\tilde{E})^{-1}\{S^{\tilde{E}}(X) + \tau^{\tilde{E}}(X)E\} \\ &= -d\varphi(\tilde{E})^{-1} S^{\tilde{E}}(X). \end{aligned}$$

Hence $S^E = -d\varphi(\tilde{E})^{-1} S^{\tilde{E}}$ and $\tau^E = 0$ hold. □

It is known that the structure induced by a non-degenerate equiaffine immersion is the statistical manifold structure. Conversely, Kurose proved the next proposition.

Proposition 3.5. ([K2]) *A simply connected statistical manifold can be realized in $\mathbf{A}^{n+1}$ by a non-degenerate equiaffine immersion if and only if it is 1-conformally flat. Such an immersion is uniquely determined up to affine transformations of $\mathbf{A}^{n+1}$*

Proposition 3.5 can be proved by projective flatness of the dual connection of a given connection ([DNV]). Finally, let us show Theorem 3.1.

Proof of Theorem 3.1. By Lemma 3.4 and Proposition 3.5 a statistical manifold (M, D^E, g^E) is 1-conformally flat. Thus Theorem 3.1 holds by Lemma 3.3. □

3.1.3 Realization of 1-conformal flatness

Let (N, ∇, h) be a 1-conformally flat statistical manifold of $\dim n \geq 2$ with a Riemannian metric h. By Proposition 3.5 (N, ∇, h) can be realized by a non-degenerate equiaffine immersion. We denote by (x, E) a non-degenerate equiaffine immersion into $\mathbf{A}^{n+1}$ which realizes (N, ∇, h). Then we can immerse (N, ∇, h) into a flat statistical manifold as the next lemma.

Lemma 3.6. *For a simply connected open subset U of N and a small $\varepsilon > 0$, we define a function φ on $\tilde{U} = \bigcup_{q \in U}\{x(q) \oplus (-\varepsilon, \varepsilon) \cdot E_q\}$ by*

$$\varphi(p) = e^{-t} \text{ for } p = x(p_0) + tE_{p_0},\ p_0 \in U, t \in (-\varepsilon, \varepsilon).$$

Then (U, ∇, h) is a statistical submanifold of a flat statistical manifold $(\tilde{U}, \tilde{D}, \tilde{D}d\varphi)$.

Proof. For $X, Y \in \mathcal{X}(U)$, we have

$$d\varphi(X) = 0, \quad d\varphi(E) = -1,$$

and

$$\begin{aligned}(\tilde{D}_X d\varphi)(Y) &= X(d\varphi(Y)) - d\varphi(\tilde{D}_X Y)\\ &= -d\varphi(\nabla_X Y + h(X, Y)E)\\ &= -h(X, Y)d\varphi(E)\\ &= h(X, Y).\end{aligned}$$

Thus, (U, ∇, h) is a submanifold of $(\tilde{U}, \tilde{D}, \tilde{D}d\varphi)$.

We also denote by E a vector field on $\tilde{U}$ whose value is E_{p_0} at $p = x(p_0) + tE_{p_0}$. On $x(U)$ we have

$$E(d\varphi(E)) = 1, \quad \tilde{D}_E E = 0,$$

and

$$(\tilde{D}_E d\varphi)(E) = E(d\varphi(E)) - d\varphi(\tilde{D}_E E) = 1.$$

Thus $(\tilde{D}d\varphi)_{x(p_0)}$ is positive definite for $p_0 \in U$. From continuity of a function φ, $\tilde{D}d\varphi$ is a Riemannian metric on $\tilde{U}$ for a small ε. Hence $(\tilde{U}, \tilde{D}, \tilde{D}d\varphi)$ is a flat statistical manifold. □

3.2 α-conformal equivalence of level surfaces

In this section, we treat α-connections of flat statistical manifolds. We consider α-conformal equivalence of induced connections on level surfaces by the α-connections. Moreover we give a procedure to realize a statistical manifold, which is α-conformally equivalent to a manifold with an α-connection, as a statistical submanifold.

3.2.1 α-connections on statistical manifolds

Let N be a manifold with a dually flat structure (∇, ∇', h). For a real number α, an affine connection defined by

$$\nabla^{(\alpha)} := \frac{1+\alpha}{2}\nabla + \frac{1-\alpha}{2}\nabla'$$

is called an α-connection of (N, ∇, h). The triple $(N, \nabla^{(\alpha)}, h)$ is also a statistical manifold, and $\nabla^{(-\alpha)}$ the dual connection of $\nabla^{(\alpha)}$. The 1-connection, the (-1)-connection and the 0-connection coincide with ∇, ∇' and Levi-Civita connection of (N, h), respectively. An α-connection is not always flat ([A]).

3.2.2 α-connections and α-conformal equivalence

We relate an α-connection of a flat statistical manifold with an α-conformal equivalence of its statistical submanifold. The next Lemma holds.

Lemma 3.7. *Let (N, ∇, h) be a flat statistical manifold, and (M, D, g) a 1-conformally flat statistical submanifold realized in (N, ∇, h). Let M_o be a simply connected open set of M. If (M_o, D, g) is 1-conformally equivalent to a flat statistical manifold $(M_o, \bar{D}, \bar{g})$, $(M_o, D^{(\alpha)}, g)$ is α-conformally equivalent to $(M_o, \bar{D}^{(\alpha)}, \bar{g})$, where $D^{(\alpha)}$ the induced connection on M_o by an α-connection $\nabla^{(\alpha)}$ of (N, ∇, h), and $\bar{D}^{(\alpha)}$ an α-connection of $(M_o, \bar{D}, \bar{g})$.*

proof. Let D' and $\bar{D}'$ be the dual connection of D and $\bar{D}$, respectively. Since $D^{(\alpha)}$ is induced by $\nabla^{(\alpha)}$,

$$D^{(\alpha)} = \frac{1+\alpha}{2}D + \frac{1-\alpha}{2}D' \quad \text{on } M_o \tag{3.7}$$

holds. For 1-(resp. (-1)-)conformal equivalence of (D, g) and $(\bar{D}, \bar{g})$ (resp. of (D', g) and $(\bar{D}', \bar{g})$), there exists a function ϕ on M_o such that

$$\bar{g}(X, Y) = e^{\phi}g(X, Y), \tag{3.8}$$

$$g(\bar{D}_X Y, Z) = g(D_X Y, Z) - d\phi(Z)g(X, Y), \quad \text{and} \tag{3.9}$$

$$g(\bar{D}'_X Y, Z) = g(D'_X Y, Z) + d\phi(X)g(Y, Z) + d\phi(Y)g(X, Z) \tag{3.10}$$

for $X, Y, Z \in \mathcal{X}(M_o)$. From (3.9) and (3.10), it follows that

$$\begin{aligned}& g((\frac{1+\alpha}{2}\bar{D} + \frac{1-\alpha}{2}\bar{D}')_X Y, Z)\\ =\ & g((\frac{1+\alpha}{2}D + \frac{1-\alpha}{2}D')_X Y, Z) - \frac{1+\alpha}{2}d\phi(Z)g(X, Y)\\ & + \frac{1-\alpha}{2}\{d\phi(X)g(Y, Z) + d\phi(Y)g(X, Z)\}.\end{aligned}$$

By (3.7) and the definition of an α-connection of $(M_o, \bar{D}, \bar{g})$,

$$\begin{aligned} g(\bar{D}^{(\alpha)}_X Y, Z) &= g(D^{(\alpha)}_X Y, Z) - \frac{1+\alpha}{2} d\phi(Z) g(X,Y) \\ &\quad + \frac{1-\alpha}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\} \end{aligned}$$

holds. This implies Lemma 3.7. □

3.2.3 Realization of α-equivalence

We call $(N, \tilde{\nabla}, h)$ a statistical manifold with an α-connection if there exists a flat statistical manifold (N, ∇, h) such that $\tilde{\nabla}$ coincides with an α-connection of (N, ∇, h). In this section, we give a procedure to realize a statistical manifold, which is α-conformally equivalent to a statistical manifold with an α-connection, in another statistical manifold as a statistical submanifold of codimension one.

For a statistical manifold which is α-conformally equivalent to a statistical manifold with an α-connection, we obtain the next theorem.

Theorem 3.8. *A statistical manifold of* $\dim n \geq 2$ *with a Riemannian metric, which is α-conformally equivalent to a statistical manifold with an α-connection for non-zero $\alpha \in \mathbf{R}$, can be locally realized as a submanifold of a statistical manifold of* $\dim(n+1)$ *with an α-connection.*

For the proof of Theorem 3.8, we show the next lemma.

Lemma 3.9. *For non-zero $\alpha \in \mathbf{R}$, let $(M, D^{(\alpha)}, g)$ be an α-conformally equivalent statistical manifold to $(M, \bar{D}^{(\alpha)}, \bar{g})$, where $\bar{D}^{(\alpha)}$ is an α-connection of a flat statistical manifold $(M, \bar{D}, \bar{g})$. Set $D^{(\beta)} := D^{LC} + \frac{\beta}{\alpha}(D^{(\alpha)} - D^{LC})$ for an arbitrary $\beta \in \mathbf{R}$, where D^{LC} the Levi-Civita connection of (M, g). Then $(M, D^{(\beta)}, g)$ is β-conformally equivalent to $(M, \bar{D}^{(\beta)}, \bar{g})$.*

proof. First, we show that (M, D^{LC}, g) is 0-conformally equivalent to $(M, \bar{D}^{(0)}, \bar{g})$. Recall that the 0-connection $\bar{D}^{(0)}$ is the Levi-Civita connection. Setting by $D^{(\alpha)\prime}$ the dual connection of $D^{(\alpha)}$, we have $(-\alpha)$-conformal equivalence of $(M, D^{(\alpha)\prime}, g)$ and $(M, \bar{D}^{(-\alpha)}, \bar{g})$ from a fact described in Subsection 3.2.1. Thus we obtain that

$$\begin{aligned} & g(\bar{D}^{(0)}_X Y, Z) = g((\frac{1}{2}\bar{D}^{(\alpha)} + \frac{1}{2}\bar{D}^{(-\alpha)})_X Y, Z) \\ = \ & \frac{1}{2}\{g(D^{(\alpha)}_X Y, Z) - \frac{1+\alpha}{2} d\phi(Z) g(X,Y) \end{aligned}$$

$$\begin{aligned} & +\frac{1-\alpha}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\} \\ & +g(D^{(\alpha)\prime}_X Y, Z) - \frac{1-\alpha}{2} d\phi(Z) g(X,Y) \\ & +\frac{1+\alpha}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\}\} \\ = \ & g((\frac{1}{2}D^{(\alpha)} + \frac{1}{2}D^{(\alpha)\prime})_X Y, Z) - \frac{1}{2} d\phi(Z) g(X,Y) \\ & +\frac{1}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\} \\ = \ & g(D^{LC}_X Y, Z) - \frac{1}{2} d\phi(Z) g(X,Y) + \frac{1}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\} \end{aligned}$$

for a certain function ϕ on $M_o \subset M$. This implies 0-conformal equivalence of (M, D^{LC}, g) and $(M, \bar{D}^{(0)}, \bar{g})$.

By definitions of $\bar{D}^{(\alpha)}$ and $\bar{D}^{(\beta)}$, $\bar{D}^{(\beta)} = \bar{D}^{(0)} + \frac{\beta}{\alpha}(\bar{D}^{(\alpha)} - \bar{D}^{(0)})$ holds. Hence it follows that

$$\begin{aligned} & g(D^{(\beta)}_X Y, Z) = g((D^{LC} + \frac{\beta}{\alpha}(D^{(\alpha)} - D^{LC}))_X Y, Z) \\ = \ & \frac{\alpha-\beta}{\alpha} g(D^{LC}_X Y, Z) + \frac{\beta}{\alpha} g(D^{(\alpha)}_X Y, Z) \\ = \ & \frac{\alpha-\beta}{\alpha}\{g(\bar{D}^{(0)}_X Y, Z) + \frac{1}{2} d\phi(Z) g(X,Y) \\ & -\frac{1}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\}\} \\ & +\frac{\beta}{\alpha}\{g(\bar{D}^{(\alpha)}_X Y, Z) + \frac{1+\alpha}{2} d\phi(Z) g(X,Y) \\ & -\frac{1-\alpha}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\}\} \\ = \ & g((\bar{D}^{(0)} + \frac{\beta}{\alpha}(\bar{D}^{(\alpha)} - \bar{D}^{(0)}))_X Y, Z) + \frac{1+\beta}{2} d\phi(Z) g(X,Y) \\ & -\frac{1-\beta}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\} \\ = \ & g(\bar{D}^{(\beta)}_X Y, Z) + \frac{1+\beta}{2} d\phi(Z) g(X,Y) \\ & -\frac{1-\beta}{2}\{d\phi(X)g(Y,Z) + d\phi(Y)g(X,Z)\}. \end{aligned}$$

This implies Lemma 3.9. □

Finally, we shall prove Theorem 3.8.

proof of Theorem 3.8. We use same notations in Lemma 3.9. Let M be a manifold of $\dim n \geq 2$, and g, $\bar{g}$ Riemannian metrices. By Lemma 3.9, $(M, D^{(1)}, g)$ is 1-conformally equivalent to a flat statistical manifold $(M, \bar{D}^{(1)}, \bar{g})$. By Theorem 3.2, $(M, D^{(1)}, g)$ can be locally realized as a submanifold of a flat statistical manifold of $\dim(n+1)$. Suppose that $(M_o, D^{(1)}, g)$ is realized in a flat statistical manifold (N, ∇, h) for a simply connected open set $M_o \subset M$. Let $D^{(\alpha)}_{sub}$ be the induced connection on M_o by an α-connection $\nabla^{(\alpha)}$ of (N, ∇, h). By Lemma 3.7, $(M, D^{(\alpha)}_{sub}, g)$ is α-conformally equivalent to $(M, \bar{D}^{(\alpha)}, \bar{g})$. Moreover, $D^{(\alpha)}_{sub} = D^{LC} + \alpha(D^{(1)} - D^{LC})$ holds by (3.7). Considering the definition of $D^{(1)}$, we have $D^{(\alpha)} = D^{LC} + \alpha(D^{(1)} - D^{LC})$. Thus $D^{(\alpha)}_{sub}$ coincides with $D^{(\alpha)}$. Hence $(M, D^{(\alpha)}, g)$, which is α-conformally equivalent to a statistical manifold with an α-connection, can be realized in $(N, \nabla^{(\alpha)}, h)$ as a submanifold of codimension one. □

Chapter 4

Divergences of statistical manifolds

4.1 Foliations by level surfaces

In this section, we describe dual-projective flatness of an affine connection D on a level surface M and projectively flatness of the dual-connection D' of D. Moreover, we construct the orthogonal foliations by level surfaces.

4.1.1 Dual-projectively flat connections

Let (N, h) be a pseudo-Riemannian manifold. Torsion free affine connections ∇ and $\bar{\nabla}$ on N are projectively equivalent if there exists a 1-form κ such that

$$\bar{\nabla}_X Y = \nabla_X Y + \kappa(X)Y + \kappa(Y)X$$

for $X, Y \in \mathcal{X}(N)$. An affine connection ∇ is called projectively flat if ∇ is locally projectively equivalent to a flat affine connection. Torsion free affine connections ∇ and $\bar{\nabla}$ on N are dual-projectively equivalent if there exists a 1-form κ such that

$$h(\bar{\nabla}_X Y, Z) = h(\nabla_X Y, Z) - \kappa(Z)h(X, Y)$$

for $X, Y, Z \in \mathcal{X}(N)$. An affine connection ∇ is called dual-projectively flat if ∇ is locally dual-projectively equivalent to a flat affine connection ([I]).

Recall that a statistical manifold (N, ∇, h) is 1-conformally flat if and only if the dual statistical manifold (N, ∇', h) is (-1)-conformally flat. Moreover, Kurose showed that, by Proposition 9.1 in [NSi], a statistical manifold (N, ∇', h) is (-1)-conformally flat if and only if ∇' is a projectively flat connection with symmetric Ricci tensor, and that

Proposition 4.1. ([K2])*A statistical manifold (N, ∇, h) is 1-conformally flat if and only if the dual connection ∇' is a projectively flat connection with symmetric Ricci tensor.*

On projective flatness, Ivanov described the next proposition in Section 2 of [I].

Proposition 4.2. ([I])*A statistical manifold (N, ∇, h) is 1-conformally flat if and only if ∇ is a dual-projectively flat connection with symmetric Ricci tensor.*

For a level surface of a Hessian domain, we obtain the next corollary of Theorem 3.1 by Proposition 4.1 and 4.2.

Corollary 4.3. *Let M be a simply connected n-dimensional level surface of φ on an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$ with a Riemannian metric $\tilde{g}$, and suppose that $n \geq 2$. Let (M, D, g) be a statistical submanifolds of $(\Omega, \tilde{D}, \tilde{g})$ and D' the dual connection of D. Then, D is a dual-projectively flat connection with symmetric Ricci tensor and D' is a projectively flat connection with symmetric Ricci tensor.*

4.1.2 Foliations by level surfaces

We denote by $\mathcal{F}$ and $\mathcal{F}^{\perp}$ a foliation on Ω_o defined by level surfaces of φ and a foliation by the gradient flow $\tilde{E}$, respectively. In this section we relate these orthogonal foliations with the dualistic structure $(\tilde{D}, \tilde{D}', \tilde{g})$.

Let $M, \hat{M}$ be two leaves of $\mathcal{F}$, and $(M, D, g), (\hat{M}, \hat{D}, \hat{g})$ the statistical submanifolds of $(\Omega, \tilde{D}, \tilde{g})$. We denote by E the vector field on Ω_o defined by (3.1), and by $\iota, \hat{\iota}$ the restriction of $\tilde{\iota}$ to $M, \hat{M}$, respectively. Non-degenerate affine immersions $(x, E), (\hat{x}, E)$ realize $(M, D, g), (\hat{M}, \hat{D}, \hat{g})$ in $\mathbf{A}^{n+1}$, where $x, \hat{x}$ are canonical immersions of $M, \hat{M}$ into Ω, respectively.

Then ι is said to be the conormal immersion for x. In fact, denoting by $\langle a, b\rangle$ a pairing of $a \in \mathbf{A}^*_{n+1}$ and $b \in \mathbf{A}^{n+1}$, we have

$$\langle \iota(p), Y_p\rangle = 0 \text{ for } Y_p \in T_pM, \quad \langle \iota(p), E_p\rangle = 1$$

for $p \in M$, considering $T_p\mathbf{A}^{n+1}$ with $\mathbf{A}^{n+1}$. Moreover, ι satisfies that

$$\langle \iota_*(Y), E\rangle = 0, \quad \langle \iota_*(Y), X\rangle = -g(Y, X)$$

and

$$\tilde{D}^*_X \iota_*(Y) = \iota_*(D'_X Y) - g'(X, Y)\iota$$

for $X, Y \in \mathcal{X}(M)$, where D' is the dual connection of D and g' the second fundamental form. Since g is non-degenerate, an immersion $\iota : M \to \mathbf{A}^*_{n+1} - \{0\}$ is a centro-affine hypersurface. Similarly a conormal immersion $\hat{\iota} : \hat{M} \to \mathbf{A}^*_{n+1} - \{0\}$ for $\hat{x}$ is also a centro-affine hypersurface ([NP]).

We set $(e^{\lambda})(p) = e^{\lambda(p)}$ for $p \in M$ and the function λ on M such that $e^{\lambda(p)}\iota(p) \in \hat{\iota}(\hat{M})$. We define a mapping $\pi : M \to \hat{M}$ by

$$\hat{\iota} \circ \pi = e^{\lambda}\iota.$$

We denote by $\bar{D}'$ an affine connection on M defined by

$$\pi_*(\bar{D}'_X Y) = \hat{D}'_{\pi_*(X)}\pi_*(Y) \quad \text{for } X, Y \in \mathcal{X}(M),$$

and by $\bar{g}$ a Riemannian metric on M such that

$$\bar{g}(X, Y) = e^{\lambda} g(X, Y). \tag{4.1}$$

Theorem 4.4. *For affine connections $D', \bar{D}'$ on M, we have*
(i) D' and $\bar{D}'$ are projectively equivalent.
(ii) (M, D', g) and $(M, \bar{D}', \bar{g})$ are (-1)-conformally equivalent.

Proof. By definition of π, $\bar{D}'$ is the connection on M induced by $e^{\lambda}\iota$. Since D' is induced by ι, from a property of centro-affine hypersurfaces, it follows (cf.[NP]) that

$$\bar{D}'_X Y = D'_X Y + d\lambda(X)Y + d\lambda(Y)X. \tag{4.2}$$

Thus (i) holds.

Statistical manifolds (M, D', g) and $(M, \bar{D}', \bar{g})$ are by definition (-1)-conformally equivalent if they satisfy (4.1) and (4.2). Thus (ii) holds. □

We denote by $\bar{D}$ an affine connection on M defined by

$$\pi_*(\bar{D}_X Y) = \hat{D}_{\pi_*(X)}\pi_*(Y) \quad \text{for } X, Y \in \mathcal{X}(M).$$

From duality of $\hat{D}$ and $\hat{D}'$, $\bar{D}$ is the dual connection of $\bar{D}'$ on M. Then the next theorem holds.

Theorem 4.5. *For affine connections $D, \bar{D}$ on M, we have*
(i) D and $\bar{D}$ are dual-projectively equivalent.
(ii) (M, D, g) and $(M, \bar{D}, \bar{g})$ are 1-conformally equivalent.

Proof. We have

$$g(\bar{D}_X Y, Z) = g(D_X Y, Z) - d\lambda(Z) g(X, Y) \tag{4.3}$$

which is equivalent (4.2) ([K2]). Affine connections D and $\bar{D}$ are by definition dual-projectively equivalent if $g(\bar{D}_X Y, Z) = g(D_X Y, Z) - \kappa(Z) g(X, Y)$ for some 1-form κ ([I]). Thus (i) holds.

Statistical manifolds (M, D, g) and $(M, \bar{D}, \bar{g})$ are 1-conformally equivalent if they satisfy (4.1) and (4.3). Thus (ii) holds. □

For $\mathcal{F}^{\perp}$, we have:

Proposition 4.6. *Every leaf of the foliation $\mathcal{F}^{\perp}$ is a $\tilde{D}'$-geodesic on Ω_o under a certain parametrization.*

Proof. It suffices to see that any integral curve of $\tilde{E}$ is a $\tilde{D}'$-geodesic. To see it, we consider the flow

$$(\frac{dx^1}{dt}, \cdots, \frac{dx^n}{dt}) = \tilde{E}. \tag{4.4}$$

The i-th coordinate of $\tilde{E}$ is $\tilde{g}^{ij}\frac{\partial \varphi}{\partial x^j}$, where $\tilde{g}_{ij} = \tilde{g}(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$ and $(\tilde{g}^{ij})$ is the inverse matrix of $(\tilde{g}_{ij})$. Since $\tilde{g}_{ij} = \frac{\partial^2 \varphi}{\partial x^i \partial x^j}$ and $x'_i = -\frac{\partial \varphi}{\partial x^i}$, we have $\tilde{g}_{ij} = -\frac{\partial x'_i}{\partial x^j}$. Then the flow (4.4) is written by $-\tilde{g}^{ij}\frac{dx'_j}{dt} = -\tilde{g}^{ij} x'_j$, i.e.,

$$\frac{dx'_i}{dt} = x'_i.$$

Thus, for an initial point $x'(0) = \{x'_1(0), \dots, x'_{n+1}(0)\} \in \Omega_o$, the integral curve of the flow (4.4) is described by

$$x'_i(t) = e^t x'_i(0).$$

Hence the integral curve of $\tilde{E}$ is a straight line with respect to an affine coordinate $\{x'_1, \dots, x'_{n+1}\}$, and the image of the integral curve is a $\tilde{D}'$-geodesic on Ω_o under a certain parametrization. □

In [A] orthogonal foliations is constructed only by flat submanifolds, and we extended to the case of 1-conformally flat statistical submanifolds. From the proof of Proposition 4.6, we can obtain a leaf of $\mathcal{F}^{\perp}$ by a dilation of a position vector of a point in $\Omega^* = \tilde{\iota}(\Omega)$.

Corollary 4.7. *For $p \in L \in \mathcal{F}^{\perp}$ we have*

$$\tilde{\iota}(L) = \{e^t \tilde{\iota}(p) \mid t \in \mathbf{R}\} \cap \Omega^*.$$

4.2 Decomposition of divergences

We give decomposition of divergences of flat statistical manifolds by a procedure different from Nagaoka and Amari's one.

An original reason for our investigation of divergences is that divergences are canonical contrast functions, which generate statistical manifolds. On contrast functions and minimum contrast leaves, see [Eg][Mm]. Divergences of conformally-projectively flat statistical manifolds are described in [Mz]. Shima studied the Riemannian foliations on Hessian domains deeply ([S2]).

4.2.1 Divergences and orthogonal foliations

First we define divergences of statistical manifolds.

Definition 4.8. ([A]) The *divergence* ρ of a flat statistical manifold $(\Omega, \tilde{D}, \tilde{g})$ is defined by

$$\rho(p, q) = \varphi(p) + \varphi^*(\iota(q)) + \sum_i x^i(p) x'_i(q) \quad \text{for } p, q \in \Omega,$$

where φ^* is the Legendre transform of φ.

Definition 4.9. ([K2]) Let (N, ∇, h) be a 1-conformally flat statistical manifold realized by a non-degenerate affine immersion (v, ξ) into $\mathbf{A}^{n+1}$, and w the conormal immersion for v. Then the divergence ρ_{conf} of (N, ∇, h) is defined by

$$\rho_{conf}(p, q) = \langle w(q), v(p) - v(q) \rangle \quad \text{for } p, q \in N.$$

The definition of ρ_{conf} is independent of the choice of a realization of (N, ∇, h).

Let N be a manifold, and ρ a function on $N \times N$. For $X_1, \cdots, X_i, Y_1, \cdots, Y_j \in \mathcal{X}(N)$, we define a function $\rho[X_1 \cdots X_i \mid Y_1 \cdots Y_j]$ on N by

$$\rho[X_1 \cdots X_i \mid Y_1 \cdots Y_j] = (X_1, 0) \cdots (X_i, 0)(0, Y_1) \cdots (0, Y_j) \rho(p, p) \quad \text{for } p \in N.$$

We call ρ a contrast function if

(i) ρ vanishes on the diagonalset of $N \times N$,

(ii) $\rho[X \mid\,] = \rho[\, \mid X] = 0$, and

(iii) h is a pseudo-Riemannian metric on N, set as $h(X, Y) = -\rho[X \mid Y]$, $X, Y \in \mathcal{X}(N)$.

For instance, the square of a Riemannian metric function is a contrast function ([Eg]).

It is known that an arbitrary statistical manifold is induced by a contrast function ([Mm]). These divergences are contrast functions of a flat statistical manifold and of a 1-conformally flat statistical manifold.

For $M \in \mathcal{F}$, we denote by ρ_{conf} the divergence of (M, D, g) induced by a non-degenerate equiaffine immersion (x, E) by Definition 4.9. Since (M, D, g) is a submanifold of $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$, we can define the divergence ρ_{sub} of (M, D, g) by the restriction of the divergence of $(\Omega, \tilde{D}, \tilde{g})$ defined by Definition 4.8, i.e., $\rho_{sub}(p, q) = \rho(p, q)$. Then we obtain:

Theorem 4.10. *For a 1-conformally flat statistical submanifold (M, D, g) of $(\Omega, \tilde{D}, \tilde{g})$, two divergences ρ_{conf} and ρ_{sub} coincide each other.*

For $p \in \Omega$ and $q \in M$, we set $\tilde{\rho}(p, q) = \langle \iota(q), x(p) - x(q) \rangle$, where ι is the conormal immersion for x. The function $\tilde{\rho}(p, \cdot)$ is called the affine distance function for (x, E) from p. For the proof of Theorem 4.10, we describe the divergence ρ by the affine distance function $\tilde{\rho}$.

Lemma 4.11. *we have*

$$\rho(p, q) = \varphi(p) - \varphi(q) + \tilde{\rho}(p, q) \quad \text{for } p \in \Omega, q \in M.$$

Proof. Since $\varphi^*(\iota(q)) = -\sum_i x^i(q)x'_i(q) - \varphi(q)$, it follows that

$$\rho(p, q) = \varphi(p) - \varphi(q) + \sum_i x'_i(q)(x^i(p) - x^i(q)). \tag{4.5}$$

Equations $\sum_i x'_i(q)(x^i(p) - x^i(q)) = \langle \iota(q), x(p) - x(q) \rangle = \tilde{\rho}(p, q)$ imply Lemma 4.11. □

Proof of Theorem 4.10. For $p, q \in M$, $\varphi(p) = \varphi(q)$ holds. Since $\rho_{sub}(p, q) = \rho(p, q)$ and $\rho_{conf}(p, q) = \tilde{\rho}(p, q)$, by Lemma 4.11 we have

$$\rho_{sub}(p, q) = \rho_{conf}(p, q).$$

□

Let us denote both ρ_{sub} and ρ_{conf} by the same notation ρ.

We can apply Lemma 4.11 to a point $q \in \Omega_o$. For a point $r \in \Omega$ such that $d\varphi_r = 0$, $x'_i(r) = 0$ holds, and thus we have $\varphi^*(\iota(r)) = -\varphi(r)$ by the definition of the Legendre transform. Hence we have by Definition 4.8:

Corollary 4.12. *For $p, r \in \Omega$ such that $d\varphi_r = 0$, we have*

$$\rho(p, r) = \varphi(p) - \varphi(r).$$

4.2.2 Projection by the minimum divergence

We shall describe the decomposition of the divergence of a flat statistical manifold $(\Omega, \tilde{D}, \tilde{g})$ with respect to orthogonal foliations $\mathcal{F}$ and $\mathcal{F}^\perp$.

Theorem 4.13. *Let (M, D, g) be a 1-conformally flat statistical submanifold of an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$, where M is an n-dimensional level surface of φ, and let $p, q \in M$, $r \in \Omega$. If a tangent vector at q, of the $\tilde{D}'$-geodesic through q and r, is vertical to T_qM with respect to $\tilde{g}$, we have*

$$\rho(p, r) = \mu\rho(p, q) + \rho(q, r), \tag{4.6}$$

where $\tilde{\iota}$ is the gradient mapping of φ defined by (1) in section 2 and $\tilde{\iota}(r) = \mu\tilde{\iota}(q)$, $\mu \in \mathbf{R}$.

Proof. Recall that ι is the restriction of $\tilde{\iota}$ to M, and using $x'_i = x^*_i \circ \tilde{\iota}$ and Definition 4.8, we have

$$\begin{aligned} & \rho(p, q) + \rho(q, r) \\ = \;& \varphi(p) - \varphi(r) + \sum_i (x'_i(r) - x'_i(q))(x^i(q) - x^i(p)) + \sum_i x'_i(r)(x^i(p) - x^i(r)) \\ = \;& \varphi(p) - \varphi(r) + \langle \tilde{\iota}(r) - \tilde{\iota}(q), x(q) - x(p) \rangle + \langle \tilde{\iota}(r), x(p) - x(r) \rangle. \end{aligned}$$

By Lemma 4.11, $\rho(p, r) = \varphi(p) - \varphi(r) + \langle \tilde{\iota}(r), x(p) - x(r) \rangle$ holds. Thus we get

$$\rho(p, r) = \rho(p, q) + \rho(q, r) + \langle \tilde{\iota}(r) - \tilde{\iota}(q), x(p) - x(q) \rangle.$$

From Corollary 4.7 the trajectory $C \in \Omega$ of the $\tilde{D}'$-geodesic, through q, r and vertical to T_qM, satisfies that

$$\{e^t\tilde{\iota}(q) \mid t \in \mathbf{R}\} \cap \Omega^* \subset \tilde{\iota}(C).$$

Thus there exists a real number μ such that $\tilde{\iota}(r) = \mu\tilde{\iota}(q)$. Since $\rho(p, q) = \langle \tilde{\iota}(q), x(p) - x(q) \rangle$, we obtain

$$\langle \tilde{\iota}(r) - \tilde{\iota}(q), x(p) - x(q) \rangle = (\mu - 1)\rho(p, q).$$

Thus, we obtain Theorem 4.13. □

By this decomposition we can obtain the projection of a point in Ω_o to $M \in \mathcal{F}$ along a leaf of $\mathcal{F}^\perp$.

Corollary 4.14. *Let M be an arbitrary leaf of $\mathcal{F}$ and r in $\Omega_o = \{p \in \Omega \mid d\varphi_p \neq 0\}$. Then the unique minimizer of a function $\rho(\cdot, r)$ on M is the intersection point of L_r and M, where L_r is the leaf of $\mathcal{F}^\perp$ including r.*

Proof. Let q be the intersection of L_r and M. Since both q and r are in L_r, there exists a positive number μ which satisfies (4.6). From positivity of divergences the point q is the unique minimizer of a function $\rho(\cdot, r)$. □

We denote by ρ' the divergence of the dual statistical manifold $(\Omega, \tilde{D}', \tilde{g})$ of $(\Omega, \tilde{D}, \tilde{g})$. Then $\rho(p, q) = \rho'(q, p)$ holds. Therefore, on the same assumption of Theorem 4.13, it follows that

$$\rho'(r, p) = \rho'(r, q) + \mu\rho'(q, p).$$

Recalling that divergences are contrast functions, and virtue of Corollary 4.14, we can call leaves of $\mathcal{F}^\perp$ minimum contrast leaves with respect to the dual divergence ρ' ([Eg]).

4.3 Gradient flow and divergences

We give examples of the gradient flow along geodesics relative to the dual connection.

On dynamical systems constrained to flat submanifolds, Fujiwara and Amari showed the following theorem and its applications to engineering.

Theorem 4.15.([FA,Theorem 2]) *Let $N = \{p_\xi \mid \xi \in \Xi \subset \mathbf{R}^n\}$ be a submanifold embedded in a flat manifold $\tilde{N}$ with respect to a dualistic structure $(\tilde{\nabla}, \tilde{\nabla}', \tilde{h})$, and (∇, ∇', h) the induced dualistic structure on N. If N is $\tilde{\nabla}$-autoparallel, then for $r \in \tilde{N}$ the gradient flow*

$$\frac{d\xi^i}{dt} = -h^{ij}\frac{\partial}{\partial \xi^j}\rho(p_\xi, r)$$

converges to a unique stationary point independent of the initial point along a ∇'-geodesic, where $\xi = (\xi^1, \cdots, \xi^n)$ is a ∇-affine coordinate such that $\nabla_X \frac{\partial}{\partial \xi^j} = 0$ for $X \in \mathcal{X}(N)$, $h_{ij} = h(\frac{\partial}{\partial \xi^i}, \frac{\partial}{\partial \xi^j})$, $[h^{ij}] = [h_{ij}]^{-1}$, and ρ is the divergence of $(\tilde{N}, \tilde{\nabla}, \tilde{h})$. Then the stationary point $q \in M$ is the unique one such that

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

A $\tilde{\nabla}$-autoparallel statistical submanifold of a flat statistical manifold $(\tilde{N}, \tilde{\nabla}, \tilde{h})$ is flat, and its divergence coincides with the restriction of the divergence of $(\tilde{N}, \tilde{\nabla}, \tilde{h})$ ([A]). These facts imply Theorem 4.15.

We shall investigate a dynamical system constrained to a 1-conformally flat statistical submanifolds. Let (M, D, g) be a 1-conformally flat statistical submanifold of an $(n+1)$-dimensional Hessian domain $(\Omega, \tilde{D}, \tilde{g} = \tilde{D}d\varphi)$, where M is an n-dimensional level surface of φ. For an affine coordinate system $\{x^1, \ldots, x^{n+1}\}$ on Ω, we consider $\{x^1, \ldots, x^n\}$ an affine coordinate system on $M_+ = \{p \in M \mid x'_{n+1}(p) > 0\}$. Let r be a fixed point such that $x'_i(r) = 0$ for $i = 1, \cdots, n$ and $x'_{n+1}(r) > 0$. We consider $\rho(\cdot, r)$ as a function on M_+ of variables $x^1, \cdots, x^n$ and denote by $\rho(p_x, r)$ its value at $p \in M_+$, where ρ is the divergence of $(\Omega, \tilde{D}, \tilde{g})$. We set $g_{ij} = g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$ for $i, j = 1, \cdots, n$ and $(g^{ij}) = (g_{ij})^{-1}$ on M_+, and then we obtain:

Proposition 4.16. *The gradient flow*

$$\frac{dx^i}{dt} = -g^{ij}\frac{\partial}{\partial x^j}\rho(p_x, r) \tag{4.7}$$

converges to the intersection of L_r and M, independent of the initial point following a D'-geodesic, where L_r is the leaf of $\mathcal{F}^\perp$ including r.

Proof. Let q be the intersection of L_r and M, and μ the positive number such that $\tilde{\iota}(r) = \mu\tilde{\iota}(q)$. By Theorem 4.13, $\rho(p, r) = \mu\rho(p, q) + \rho(q, r)$ holds. Thus the gradient flow (4.7) is equivalent to

$$\frac{dx^i}{dt} = -\mu g^{ij}\frac{\partial}{\partial x^j}\rho(p_x, q). \tag{4.8}$$

Since $\rho(p_x, q) = \sum_{i=1}^{n+1} x'_i(q)(x^i(p) - x^i(q))$ and $x'_i(q) = 0$ for $i = 1, \cdots, n$, we have

$$\frac{\partial}{\partial x^j}\rho(p_x, q) = x'_{n+1}(q)\frac{\partial x^{n+1}}{\partial x^j}, \tag{4.9}$$

considering x^{n+1} a function of variables $x^1, \cdots, x^n$. The Riemannian metric g is the fundamental form for the affine immersion (x, E) which realize

(M, D, g). Since ι is a conormal immersion for x, we have

$$g_{ij} = \langle \iota(\cdot), \tilde{D}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \rangle \quad \text{for} \quad \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \in \mathcal{X}(M_+).$$

Using $x'_i = x^*_i \circ \tilde{\iota}$ and $\tilde{D}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = (0, \cdots, 0, \frac{\partial^2 x^{n+1}}{\partial x^i \partial x^j})$, we obtain

$$g_{ij} = x'_{n+1} \frac{\partial^2 x^{n+1}}{\partial x^i \partial x^j}. \tag{4.10}$$

From (4.9) and (4.10), the flow (4.8) is described by

$$x'_{n+1} \frac{\partial^2 x^{n+1}}{\partial x^i \partial x^j} \frac{dx^i}{dt} = -\mu x'_{n+1}(q) \frac{\partial x^{n+1}}{\partial x^j}.$$

Setting $X_i = \frac{\partial x^{n+1}}{\partial x^i}$, we have

$$\frac{dX_i}{dt} = -\mu x'_{n+1}(q)(x'_{n+1})^{-1} X_i. \tag{4.11}$$

Since the function x'_{n+1} converges to $x'_{n+1}(q)$ as $x^i \to 0$, i.e., $X_i \to 0$ for $i = 1, \cdots, n$ and $x'_{n+1}(p) > 0$, X_i is monotone decreasing as $t \to +\infty$. Thus the flow (4.11) is asymptotic to the flow $\frac{dX_i}{dt} = -\mu X_i$, i.e., $X_i = Ce^{-\mu t}$, where C is a constant number, as $t \to +\infty$. Hence the flow (4.11) converges to q, which satisfies that $x'_i(q) = 0$ for $i = 1, \cdots, n$, along a straight line with respect to a coordinate system $\{X_1, \ldots, X_n\}$. From a property of normal vectors of level surfaces of φ and x^{n+1}, there exists a positive number ν_p for $p \in M_+$ such that

$$\nu_p(\frac{\partial \varphi}{\partial x^1}, \cdots, \frac{\partial \varphi}{\partial x^n})_p = (\frac{\partial x^{n+1}}{\partial x^1}, \cdots, \frac{\partial x^{n+1}}{\partial x^n})_p.$$

Then we have

$$-\nu_p(x'_1, \cdots, x'_n)_p = (\frac{\partial x^{n+1}}{\partial x^1}, \cdots, \frac{\partial x^{n+1}}{\partial x^n})_p.$$

Thus the flow (4.11) converges to q along a straight line with respect to a coordinate system $\{x'_1, \cdots, x'_n\}$. In other words, the trajectory of the flow is on the intersection of M and a plane including the origin with respect to a coordinate system $\{x'_1, \cdots, x'_n\}$. It is known that the intersection of a projectively flat surface and a plane including the center for the projective transform is a pseudo-geodesics of the surface ([Ei]). The manifold (M, D', g)

is projectively flat by Corollary 4.3. The center for the projective transform is the origin with respect to a coordinate system $\{x'_1, \cdots, x'_n\}$ by the proof of Theorem 4.4. Thus the trajectory of the flow (4.11) becomes a geodesics with respect to D'. Hence the flow (4.7) converges to q independent of the initial point following a D'-geodesic. □

Chapter 5

Dualistic strustures on symmetric cones

5.1 Jordan algebras and dual connections

In this section, we relate dual connections on symmetric cones with Jordan algebras. First we give foundations of Jordan algebras and symmetric cones.

5.1.1 Jordan algebras

A vector space V is called a Jordan algebra if a product $*$ defined on V satisfies

$$\begin{aligned} x * y &= y * x, \\ x * (x^2 * y) &= x^2 * (x * y) \end{aligned}$$

for all $x, y \in V$ by setting $x^2 = x * x$. Let V be an n-dimensional Jordan algebra over $\mathbf{R}$ with an identity element e, i.e., $x*e = e*x = x$. Denoting by $m(x)$ the degree of the minimal polynomial of $x \in V$, the rank of V is defined by $r = \max\{m(x) \mid x \in V\}$. An element $x \in V$ is said to be invertible if there exists $y \in \mathbf{R}[x]$ such that $x * y = e$, where $\mathbf{R}[X]$ is polynomials of X over $\mathbf{R}$. Since $\mathbf{R}[x]$ is an associative algebra, y is unique, called the inverse of x and denoted by $x^{-1} = y$.

For x in V, let $L(x)$ and $P(x)$ be endomorphisms of V defined by

$$\begin{aligned} L(x)y &= x * y, \ y \in V \\ P(x) &= 2L(x)^2 - L(x^2). \end{aligned}$$

Following results, about P the quadratic representation of V, are known.

Proposition 5.1. ([FK]) (i) *An element x is invertible if and only if $P(x)$ is invertible, and*

$$P(x)x^{-1} = x,$$
$$P(x)^{-1} = P(x^{-1}).$$

(ii) *If x and y are invertible, so is $P(x)y$ and*

$$(P(x)y)^{-1} = P(x^{-1})y^{-1}.$$

(iii) *For all x and y,*

$$P(P(y)x) = P(y)P(x)P(y).$$

For each $x \in V$ we identify a tangent space T_xV with V as the following

$$T_xV \ni \sum_i u^i(\frac{\partial}{\partial x^i})_x \sim u \in V, \quad x^i(u) = u^i \in \mathbf{R}$$

where $\{x^1, \dots, x^n\}$ is the canonical coordinate on V. Let $\mathcal{D}x^{-1}$ be the differential at x of a map $y \mapsto y^{-1}$ on $\mathcal{I} = \{y \in V \mid y : invertible\}$. Using the identification, we consider $\mathcal{D}x^{-1}$ maps $u \in V \simeq T_xV \simeq T_x\mathcal{I}$ to $\mathcal{D}_u x^{-1} \in V$. Let $\mathcal{D}P(x)$ and $\mathcal{D}P(x)^{-1}$ be the differentials at x of maps $y \in V \mapsto P(y) \in End(V)$ and $y \in \mathcal{I} \mapsto P(y)^{-1} \in End(V)$, respectively. In the same way, we consider $\mathcal{D}P(x)$, $\mathcal{D}P(x)^{-1}$ map $u \in V$ to $\mathcal{D}_uP(x)$, $\mathcal{D}_uP(x)^{-1} \in End(V)$, respectively.

Proposition 5.2. *The differentials are given by*
(i) $\mathcal{D}_u x^{-1} = -P(x)^{-1}u$,
(ii) $\mathcal{D}_uP(x) = P(x+u) - P(x) - P(u) = 2P(x,u)$,
where $P(x,u) = L(x)L(u) + L(u)L(x) - L(x*u)$,
(iii) $\mathcal{D}_uP(x)^{-1} = -P(x)^{-1}(\mathcal{D}_uP(x))P(x)^{-1}$.

Proof. The proof of (i),(ii) are given in [FK]. Differentiating both sides of $P(x)P(x)^{-1} = Id.$($Id.$ is an identity map), we obtain

$$(\mathcal{D}_uP(x))P(x)^{-1} + P(x)\mathcal{D}_uP(x)^{-1} = 0.$$

Thus, (iii) follows. □

Proposition 5.3. ([FK]) *The next follows:*

$$P(P(x)u, P(x)v) = P(x)P(u,v)P(x).$$

5.1.2 Symmetric cones associated with Jordan algebras

Let Ω be a open convex cone on a vector space V. We denote by G the identity component of the linear automorphism group of Ω. If G acts on Ω transitively, Ω is said to be homogeneous. The dual cone of Ω is defined by

$$\Omega^* = \{y \in V^* \mid (x,y) > 0, \forall x \in \bar{\Omega}\backslash\{0\}\},$$

where V^* is the dual vector space of V, $(\ ,\ )$ the pairing on V, $\bar{\Omega}$ the closure of Ω. If $\Omega = \Omega^*$ by identification with V and V^*, Ω is said to be self-dual. A cone Ω is called symmetric if it is homogeneous and self-dual.

A Jordan algebra V is said to be simple if there does not exist non-trivial ideal on it, and to be Euclidean (or formally real) if V has a positive definite symmetric associative bilinear form. Here a bilinear form $\langle\ ,\ \rangle$ is called associative if $\langle L(x)y, z\rangle = \langle y, L(x)z\rangle$ for all $x, y, z \in V$.

Through this section, let V be a simple Euclidean Jordan algebra and Ω be the symmetric cone associated with V, i.e., $\Omega = \{x^2 \mid x \in \mathcal{I}\}$.

Let D be the canonical flat affine connection on V. We denote by $\{x_1^*, \dots, x_n^*\}$ and D^* the dual affine coordinate system and the canonical flat affine connection on V^*, respectively. The characteristic function of Ω is defined by

$$\varphi(x) = \int_{\Omega^*} e^{-(x,y)} dy$$

where dy is the Euclidean measure on V^*. For $\psi = \frac{r}{n}\log\varphi$, we define the gradient mapping ι from Ω to V^* by

$$x_i^* \circ \iota = -\frac{\partial\psi}{\partial x^i}.$$

The image of ι corresponds with $\Omega^* \simeq \Omega$. Thus ι is the restriction of the map $x \mapsto x^{-1}$ on Ω, and the differential ι_* is described by $-P(x)^{-1} = \mathcal{D}x^{-1}$ at $x \in \Omega$ from Proposition 5.2(i). A Riemannian metric g, defined on Ω by

$$g_x(u,v) = -((\iota_*)_x u, v) = (P(x)^{-1}u, v)$$

for $x \in \Omega, u, v \in V \simeq T_xV$ (where $(\ ,\ )$ is the canonical inner product on V), is positive definite and G-invariant. The map ι is an involutive isometry on Ω with respect to g, which has unique fixed point e.

We define a flat affine connection D' on Ω by

$$\iota_*(D'_XY) = D^*_{\iota_*(X)}\iota_*(Y)$$

for $X, Y \in \mathcal{X}(\Omega)$ where $\mathcal{X}(\Omega)$ is the set of tangent vector fields on Ω. Here we consider that ι maps Ω to $\Omega^* \subset V^*$ and that $D^*_{\iota_*(X)}\iota_*(Y)$ is the covariant derivative along ι induced by D^*.

Remark 5.4. The connection D, D' are torsion-free, and the next follows:

$$Xg(Y,Z) = g(D_X Y, Z) + g(Y, D'_X Z), \quad X, Y, Z \in \mathcal{X}(\Omega).$$

Then the triple (g, D, D') is called a dualistic structure. In addition,

$$D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0, \quad D'_{\frac{\partial}{\partial x'_i}} \frac{\partial}{\partial x'_j} = 0$$

holds where $x'_i = x^*_i \circ \iota = -\dfrac{\partial \psi}{\partial x^i}$. Thus $\{x'_1, \ldots, x'_n\}$ is an affine coordinate with respect to D'. Hence D, D' are dually flat affine connections with respect to g. The dually flat structure has been studied also in terms of statistical application [A].

Remark 5.5. The Riemannian metric g is also given by $g = \sum_{i,j} \dfrac{\partial^2 \psi}{\partial x^i \partial x^j} dx^i dx^j$. Then the pair (D, g) is called Hessian structure and (Ω, D, g) is a Hessian manifold ([S1]).

We represent the connection D', using the quadratic representation P. Through this paper, we do not distinguish $T_x\Omega$ from V.

Lemma 5.6. *Setting* $u = (\dfrac{\partial}{\partial x^i})_x$, $v = (\dfrac{\partial}{\partial x^j})_x \in V$, *we have*

$$(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = P(x)\mathcal{D}_u P(x)^{-1} v.$$

Proof. Since $\iota_*(\dfrac{\partial}{\partial x'_i}) = \dfrac{\partial}{\partial x^*_i}$ and $\dfrac{\partial}{\partial x^i} = -\sum_k g_{ik} \dfrac{\partial}{\partial x'_k}$, where $g_{ij} = g(\dfrac{\partial}{\partial x^i}, \dfrac{\partial}{\partial x^j})$, we have $\iota_*(\dfrac{\partial}{\partial x^i}) = -\sum_k (g_{ik} \circ \iota^{-1}) \dfrac{\partial}{\partial x^*_k}$. Thus it follows

$$\begin{aligned} D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} &= \iota_*^{-1}(D^*_{\iota_*(\frac{\partial}{\partial x^i})} \iota_*(\frac{\partial}{\partial x^j})) \\ &= \iota_*^{-1}(D^*_{\sum_k (g_{ik}\circ\iota^{-1})\frac{\partial}{\partial x^*_k}} \sum_l (g_{jl} \circ \iota^{-1}) \frac{\partial}{\partial x^*_l}) \end{aligned}$$

$$\begin{aligned} &= \iota_*^{-1}(\sum_{k,l} (g_{ik} \circ \iota^{-1}) \frac{\partial (g_{jl} \circ \iota^{-1})}{\partial x^*_k} \frac{\partial}{\partial x^*_l}) \\ &= \sum_{k,l,r,s} g_{ik} \frac{\partial g_{jl}}{\partial x^r} g^{rk} g^{sl} \frac{\partial}{\partial x^s} \\ &= \sum_{l,s} g^{sl} \frac{\partial g_{lj}}{\partial x^i} \frac{\partial}{\partial x^s}. \end{aligned}$$

Recall $(g_{ij})_x = (P(x)^{-1}u, v)$. Setting $w = (\dfrac{\partial}{\partial x^s})_x$, we have

$$(g^{sl} \frac{\partial g_{lj}}{\partial x^i})_x = (P(x)\mathcal{D}_u P(x)^{-1} w, v) = (P(x)\mathcal{D}_u P(x)^{-1} v, w).$$

Thus, it follows

$$(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = \sum_{w \in \{(\frac{\partial}{\partial x^1})_x, \cdots, (\frac{\partial}{\partial x^n})_x\}} (P(x)\mathcal{D}_u P(x)^{-1} v, w) w = P(x)\mathcal{D}_u P(x)^{-1} v.$$

□

Corollary 5.7. *For* u, v, *we have*

$$(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -(\mathcal{D}_u P(x)) P(x)^{-1} v.$$

Proof. It follows from Proposition 5.2(iii) and Lemma 5.6. □

Let us consider relations with the group G and the quadratic representation P. In fact, $\{P(x) \mid x \in \Omega\} \subset G$ and $\Omega = \{P(y)e \mid y \in \Omega\}$ follow ([FK]). For $y \in \Omega, x = y^2$, $P(y)$ transfers an identity e to x. So $P(y)$ describes the map from $T_e\Omega$ to $T_x\Omega$ induced by the action of G. Therefore, if $a \in T_e\Omega \simeq V$ and $P(y)a \in T_x\Omega \simeq V$, we can consider a set $\bigcup_{y \in \Omega} P(y)a$ is a G-invariant vector field on Ω.

5.1.3 Dual connections and Jordan algebras

First we show an example in the case of $PD(k)$, the cone of k by k real positive definite matrices. The cone $PD(k)$ is associated with $Sym(k)$ a Jordan algebra of k by k real symmetric matrices ([FK]). So we can introduce the dually flat structure (g, D, D') on $PD(k)$ as Section 3.

It is known that a concrete description of D' at $x \in PD(k)$ is

$$(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -ux^{-1}v - vx^{-1}u$$

where we set $u = (\frac{\partial}{\partial x^i})_x$, $v = (\frac{\partial}{\partial x^j})_x \in V$ as same as the previous section and denote by uv usual matrix product with u and v ([O][OSA]). Note that the inverse as Jordan algebra coincide with usual matrix inverse ([FK]). Since $2L(u)v = uv + vu$ and $P(y)u = yuy$ on $Sym(k)$, we obtain

$$(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -2P(y)((P(y)^{-1}u) * (P(y)^{-1}v)), \quad y^2 = x.$$

In general, we have

Theorem 5.8. *On a cone (Ω, g, D, D') associated with a simple Euclidean Jordan algebra, the value of the connection D', which is pull-back from x to the identity e by the action of G, is equivalent to -2 times Jordan product of pull-back vectors, that is,*

$$P(y)^{-1}(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -2(P(y)^{-1}u) * (P(y)^{-1}v), \quad y^2 = x.$$

Proof. From Proposition 5.1(iii) we have $P(y^2) = P(y)^2$ or $P(x) = P(y)^2$, and $P(x)^{-1} = P(y)^{-2}$. By Corollary 5.7 it follows

$$P(y)^{-1}(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -P(y)^{-1}(\mathcal{D}_u P(x))P(y)^{-2}v.$$

Therefore we have only to show

$$P(y)^{-1}(\mathcal{D}_u P(x))P(y)^{-1} = 2L(P(y)^{-1}u).$$

By Proposition 5.1(i)(iii) and Proposition 5.2(ii), we have

$$\begin{aligned}
& P(y)^{-1}(\mathcal{D}_u P(x))P(y)^{-1} \\
= \; & P(y)^{-1}(P(x+u) - P(x) - P(u))P(y)^{-1} \\
= \; & P(P(y^{-1})(x+u)) - P(P(y^{-1})x) - P(P(y^{-1})u) \\
= \; & P(e + P(y^{-1})u) - P(e) - P(P(y^{-1})u) \\
= \; & 2(L(e)L(P(y^{-1})u) + L(P(y^{-1})u)L(e) - L(e * P(y^{-1})u)) \\
= \; & 2L(P(y^{-1})u).
\end{aligned}$$

Thus the theorem is proved (The similar technique has been used in [F]). □

Further, considering a *mutation* of V, we can correspond a connection with a Jordan product directly.

Definition. For f in a Jordan algebra V, we give a product $\perp_f$ by

$$u \perp_f v = P(u, v)f, \quad u, v \in V.$$

Then $(V, \perp_f)$ is a Jordan algebra, denoted by V_f and called a *mutation* of V with respect to f ([BK]).

Theorem 5.9. *The value of the connection D' at x is equivalent to -2 times Jordan product $\perp_{x^{-1}}$ on a mutation $V_{x^{-1}}$, that is,*

$$(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -2u \perp_{x^{-1}} v.$$

Proof. From Proposition 5.2, 5.3 and Theorem 5.8, we have

$$\begin{aligned}
(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x &= -2P(y)\{(P(y)^{-1}u) * (P(y)^{-1}v)\} \\
&= -2P(y)P(P(y)^{-1}u, P(y)^{-1}v)e \\
&= -2P(y)P(P(y)^{-1}u, P(y)^{-1}v)P(y)P(y)^{-1}e \\
&= -2P(u, v)x^{-1} \\
&= -2u \perp_{x^{-1}} v.
\end{aligned}$$

□

It is known that the Levi-Civita connection ∇ on (Ω, g) is the mean of connections D and D' ([A]), i.e.,

$$\nabla_X Y = \frac{1}{2}(D_X Y + D'_X Y), \quad X, Y \in \mathcal{X}(\Omega).$$

Since $D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$, we have $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \frac{1}{2} D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$. Thus the next corollaries follow.

Corollary 5.10. *On a cone (Ω, g, D, D') associated with a simple Euclidean Jordan algebra, the value of the Levi-Civita connection ∇, which is pull-back from x to the identity e by the action of G, is equivalent to the minus of Jordan product of pull-back vectors, that is,*

$$P(y)^{-1}(\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -(P(y)^{-1}u) * (P(y)^{-1}v), \quad y^2 = x.$$

Corollary 5.11. *The value of the connection ∇ at x is equivalent to the minus of Jordan product $\perp_{x^{-1}}$ on a mutation $V_{x^{-1}}$, that is,*

$$(\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = -u \perp_{x^{-1}} v.$$

Remark 5.12. If a cone Ω is reducible, in other words, if a Euclidean Jordan algebra V is not simple, we can induce the dually flat structure (g, D, D') on Ω from the function ψ, setting

$$\psi = \sum_{i=1}^{m} \frac{r_i}{n_i} \log \varphi_i.$$

Here, we have a decomposition $V = V_1 \otimes \cdots \otimes V_m$, a direct sum of simple Euclidean Jordan algebras V_i, and denote by n_i, r_i the dimension, the rank of V_i. Let Ω_i be the cone associated with V_i (then $\Omega = \Omega_1 + \cdots + \Omega_m$, a direct product) and φ_i be the characteristic function of Ω_i. Thus the above results also hold for a reducible cone Ω.

5.2 Doubly autoparallel submanifolds

We show relation with Jordan subalgebras and doubly autoparallel submanifolds of symmetric cones.

5.2.1 On dual connections

Let V, Ω be a Euclidean Jordan algebra, a cone associated with V, respectively. For a linear subspace W in V and $p \in V$, we set $W + p = \{w + p \mid w \in W\}$ and $M = (W + p) \cap \Omega$. If M is not empty, we denote by D, D' the dual connections on submanifolds M naturally induced from the connections D, D' on Ω. By D-flatness of Ω and linearity of W, M is a D-autoparallel submanifold of Ω. Conversely a D-autoparallel submanifold is represented by $(W + p) \cap \Omega$ for some W and p.

Definition. We call M *doubly autoparallel* when M is both D- and D'-autoparallel.

It is interesting to know conditions for M to be a doubly autoparallel submanifold of Ω (On $PD(k)$ they have already been studied in [O][OSA]).

Lemma 5.13. *A submanifold M is doubly autoparallel if and only if $P(y)^{-1}W$ is a Jordan subalgebra for all $x \in M$ where $y^2 = x, y \in \Omega$*

Proof. A submanifold M is D'-autoparallel if and only if $D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \in \mathcal{X}(M)$ where $\mathcal{X}(M)$ is the set of tangent vector fields on M ([A]). We denote

by $\{x^1, \ldots, x^m\}$ a D-flat affine coordinate on M. Identifying W with T_xM for $x \in M$ as Lemma 5.6, this condition equals to $(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x \in W$ or $P(y)^{-1}(D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x \in P(y)^{-1}W$. Set $u = (\frac{\partial}{\partial x^i})_x, v = (\frac{\partial}{\partial x^j})_x \in W$. By Theorem 5.8, it is equivalent to

$$(P(y)^{-1}u) * (P(y)^{-1}v) \in P(y)^{-1}W.$$

for all $u, v \in W$. Thus the lemma holds. □

When a certain condition holds, a submanifold M is doubly autoparallel if and only if its tangent space W is a Jordan subalgebra.

Theorem 5.14. *Let $M = W \cap \Omega$, i.e., $p = 0$ and suppose M includes e the identity element of V. A submanifold M is doubly autoparallel if and only if W is a Jordan subalgebra of V (Then M is a subcone in Ω associated with W).*

Proof. By Lemma 5.13, we have only to show that $P(y)^{-1}W$ is a Jordan subalgebra for all $x \in M$ if and only if W is a Jordan subalgebra.

Let $P(y)^{-1}W$ be a Jordan subalgebra for $x \in M$. Since $e \in M$ and $P(e)$ is an identity map on Ω, $P(e)^{-1}W = W$ is a Jordan subalgebra. Conversely let W be a Jordan subalgebra. Then the manifold M is the cone associated with W. The identity component of the linear automorphism group of M is described by $\{P(y) \mid y \in M\}$. Thus for $x \in M$ there exists $y \in M$ such that $P(y)^{-1}W = W, y^2 = x$. Hence $P(y)^{-1}W$ is a Jordan subalgebra. □

We shall note concerning the Levi-Civita connection ∇. The next corollaries follow by Corollary 5.10, Lemma 5.13.

Corollary 5.15. *A submanifold M is ∇-autoparallel if and only if $P(y)^{-1}W$ is a Jordan subalgebra for all $x \in M$ where $y^2 = x, y \in \Omega$*

Corollary 5.16. *Let $M = W \cap \Omega$, i.e., $p = 0$ and suppose M includes e the identity element of V. A submanifold M is ∇-autoparallel if and only if W is a Jordan subalgebra of V.*

Remark 5.17. The connections D, D', ∇ are torsion free. Thus a submanifold M is autoparallel if and only if M is totally geodesic for D, D', ∇, respectively ([KN]).

Remark 5.18. There exist doubly autoparallel submanifolds of Ω such that they do not include an identity element of V or $p \neq 0$. For example, setting $V = Sym(3), \Omega = PD(3)$,

$$W = \left\{ \begin{pmatrix} t & 0 & 0 \\ 0 & 2t & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid t \in \mathbf{R} \right\}, \; W' = \left\{ \begin{pmatrix} t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 0 \end{pmatrix} \mid t \in \mathbf{R} \right\} \subset V,$$

$$p = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \in V,$$

W is not a Jordan subalgebra of V. But $P(y)^{-1}W$ ($y^2 = x \in M = (W + p) \cap \Omega$) equals to W' which is a Jordan subalgebra. Hence M is a doubly-autoparallel submanifold of Ω.

5.2.2 On α-connections

For an α-connection we see results similar to theorems in the previous subsection.

Let $D^{(\alpha)}$ be an α-connection of a cone (Ω, g, D) associated with a Euclidean Jordan algebra V. By the definition of α-connections in subsection 3.2.1, we have

$$D^{(\alpha)} = \frac{1+\alpha}{2} D + \frac{1-\alpha}{2} D'.$$

Since $D_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = 0$, we have $D^{(\alpha)}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \frac{1-\alpha}{2} D'_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$. , Then we obtain a corollary of Theorem 5.8. We set $u = (\frac{\partial}{\partial x^i})_x, v = (\frac{\partial}{\partial x^j})_x \in W$ as Subsection 5.1.3.

Corollary 5.19 *On a cone (Ω, g, D, D') associated with a simple Euclidean Jordan algebra, the value of an α-connection $D^{(\alpha)}$, which is pull-back from x to the identity e by the action of G, is equivalent to $(\alpha - 1)$ times Jordan product of pull-back vectors, that is,*

$$P(y)^{-1}(D^{(\alpha)}_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j})_x = (\alpha - 1)(P(y)^{-1}u) * (P(y)^{-1}v), \quad y^2 = x.$$

The connections $D^{(-1)}$ and $D^{(0)}$ are the dual connection D' and Levi-Civita connection ∇, respectively. Thus for $\alpha = -1$ and $\alpha = 0$, Corollary 5.19 coincides with Theorem 5.8 and Corollary 5.10, respectively. We have the next corollary of Theorem 5.14, using same notations in Subsection 5.2.1.

Corollary 5.20 *Let $M = W \cap \Omega$, i.e., $p = 0$ and suppose M includes e the identity element of V. A submanifold M is α-autoparallel if and only if W is a Jordan subalgebra of V.*

For $\alpha = 0$, Corollary 5.20 coincides with Corollary 5.16.

Chapter 6

Conclusions

We described backgrounds and motivations in Chapter 1, and in Chapter 2 gave foundamental facts on statistical manifolds to be utilized through this dissertation .

We showed 1-conformal flatness of level surfaces of a flat statistical manifold with respect to the induced connection, by affine differential geometry on Hessian domains in Chapter 3. We also mentioned to α-conformal equivalence of level surfaces with the connections induced by α-connections of flat statistical manifolds.

We studied foliations and divergences of a flat statistical manifold in Chapter 4, using results in Chapter3 essentially. We gave the decomposition of the divergence of a flat statistical manifold from orthogonal foliations: one by 1-conformally flat submanifolds and one by geodesics with respect to the dual connection. As applications, we saw the projection of a point in a flat statistical manifold to a level surface, i.e., a 1-conformally flat submanifold, given by minimization of the divergence. Next we gave a gradient system restricted on a level surface, using the divergence.

In Chapter 5, we investigated dualistic structures of symmetric cones. We related Jordan algebras and dual connections of symmetric cones. For some condition we showed that submanifolds of symmetric cones are doubly autoparallel if and only if their tangent spaces are Jordan subalgebras. We obtained the same results in concern to α-connections.

References

[A] S. Amari: Differential-Geometrical Methods in Statistics, Lecture Notes in Statistics **28**, Springer-Verlag (1985).

[AN] S. Amari and H. Nagaoka: Method of Information Geometry, Lecture Series in Applied Mathematics **12**, Iwanami (1994), in Japanese.

[BK] H. Braun und M. Koecher: Jordan-Algebren, Grundlehren der Math. Wissenschaften **128**, Springer-Verlag (1966).

[DNV] F. Dillen, K. Nomizu and L. Vrancken: Conjugate connections and Radon's theorem in affine differential geometry, Monatsh. Math. **109** (1990), 221-235.

[Eg] S. Eguchi: Geometry of minimum contrast, Hiroshima Math. J. **22** (1992), 631-647.

[Ei] L.P. Eisenhart: Non-Riemannian Geometry, Amer. Math. Soc. Colloq. Publ. **8** (1927).

[FK] J. Faraut and A. Korányi: Analysis on Symmetric Cones, Oxford (1994).

[F] L. Faybusovich: Linear systems in Jordan Algebras and primal-dual Interior-point Algorithms, J. Computational and Applied Math. **86** (1997), 149-175.

[FA] A. Fujiwara and S. Amari: Gradient systems in view of information geometry, Physica D **80** (1995), 317-327.

[HS] J.H. Hao and H. Shima: Level surfaces of non-degenerate functions in $\mathbf{R}^{n+1}$, Geom. Dedicata **50** (1994), 193-204.

[I] S. Ivanov: On dual-projectively flat affine connections, J. of Geom. **53** (1995), 89-99.

[KN] S. Kobayashi and K. Nomizu: Foundations of differential geometry, vol.I, John Wiley & Sons, New York (1963).

[K1] T. Kurose: Dual connections and affine geometry, Math. Z. **203** (1990), 115-121.

[K2] T. Kurose: On the divergence of 1-conformally flat statistical manifolds, Tôhoku Math. J. **46** (1994), 427-433.

[Mz] H. Matsuzoe: On realization of conformally-projectively flat statistical manifolds and the divergences, Hokkaido Math. J. **27** (1998), 409-421.

[Mm] T. Matumoto: Any statistical manifold has a contrast function −On the C^3-functions taking the minimum at the diagonal of the product manifold, Hiroshima Math. J. **23** (1993), 327-332.

[N] Y. Nakamura: Gradient systems associated with probability distributions, Japan J. Indust. Appl. Math. **11** (1994), 21-30.

[NP] K. Nomizu and U. Pinkall: On the geometry and affine immersions, Math. Z. **195** (1987), 165-178.

[NSa] K. Nomizu and T. Sasaki: Affine Differential Geometry −Geometry of Affine Immersions−, Cambridge Univ. Press (1994).

[NSi] K. Nomizu and U. Simon: Notes on conjugate connections, Geometry and Topology of Submanifolds **IV** (F.Dillen and L.Verstraelen eds.), World Scientific (1992), 152-173.

[O] A. Ohara: Information Geometric Analysis of an Interior Point Method for Semidefinite Programming, O. Barndorf-Nielsen and V. Jensen eds., Proc. of Geometry in Present Day Science, World Scientific (in press).

[OA] A. Ohara and S. Amari: Differential Geometric Structures of Stable State Feedback Systems with Dual Connections, Kybernetika **30** (1984), 369-386.

[OSA] A. Ohara, N. Suda and S. Amari: Dualistic differential geometry of positive definite matrices and its applications to related problems, Linear Algebra and Its Applications **247** (1996), 31-53.

[OAT] I. Okamoto, S. Amari and K. Takeuchi: Asymptotic theory of sequential estimation: differential geometrical approach, Ann. Statist. **19** (1991), 961-981.

[S1] H. Shima: Symmetric Spaces with Invariant Locally Hessian Structures, J.Math.Soc.Japan **29** (1977), 581-589.

[S2] H.Shima: Geometry of Hessian domains, Differential Geometry and its Applications **7** (1997), 277-290.

[S3] H. Shima: Homogeneous Spaces with Invariant Projectively Flat Affine Connections, preprint.

[SH] H. Shima and J.H. Hao: Geometry associated with normal distributions, Preprint.

[UO] K. Uohashi and A. Ohara: Jordan Algebras and Dual Affine Connections on Symmetric Cones, Preprint.

[UOF1] K. Uohashi, A. Ohara and T. Fujii: 1-conformally flat statistical submanifolds, to appear in Osaka J. math.

[UOF2] K. Uohashi, A. Ohara and T. Fujii: Foliations and divergences of flat statistical manifolds, Preprint.

[Vi] E.B. Vinberg: Homogeneous convex cones, Trans. Moscow Math. Soc. **12** (1963), 340-403.

[Vo] P.W. Vos: Fundamental equations for statistical submanifolds with applications to the Bartlett correction, Ann. Inst. Statist. Math. **41** (1989), 429-450.