SUT Journal of Mathematics
Vol. 46, No. 1 (2010), 1–21

Exponential families admitting almost complex structures

Kazuhiko Takano

(Received June 5, 2009; Revised April 8, 2010)

Abstract. We discuss exponential families admitting almost complex structures which are parallel relative to an exponential connection (e–connection) or mixture connection (m–connection). The multinomial distribution, negative multinomial distribution and multivariate normal distribution are important examples of the exponential family. We give almost complex structures which are parallel relative to the exponential or mixture connection for these exponential families. Also, we prove spaces of the multinomial distribution and negative multinomial distribution are of constant curvature with respect to the α–connection.

AMS 2000 *Mathematics Subject Classification.* 53C25, 53A15, 60E05.

Key words and phrases. exponential family, α–connection, almost complex structure, statistical manifold.

§1. Introduction

Statistical models in information geometry have a Fisher metric as a Riemannian metric, and admit a torsion-free affine connection which is constructed from expectations of derivatives of a probability density ([1], [2]). This affine connection is called an α–connection, denoted by $\nabla^{(\alpha)}$, where α is a real number, and conjugate relative to the Fisher metric is a $(-\alpha)$–connection. The 0–connection is a Levi-Civita connection with respect to the Fisher metric. Particularly, $\nabla^{(1)}$ (resp. $\nabla^{(-1)}$) is said to be an exponential connection (resp. mixture connection) or e–connection (resp. m–connection) simply and denoted by $\nabla^{(e)}$ (resp. $\nabla^{(m)}$). The statistical model of an exponential family (resp. mixture family) is 1–flat (resp. (-1)–flat). The e–connection and m–connection are conjugate with respect to the Fisher metric. The e and m–connections include important concepts in information geometry.

If a density function can be expressed in terms of functions C, $F_1, \dots, F_n$ on the set χ and a function φ on Θ the subset of $\mathbb{R}^n$ as

$$p(x;\theta) = \exp\left[C(x) + \sum_{s=1}^{n} \theta^s F_s(x) - \varphi(\theta)\right], \tag{1.1}$$

then an n–dimensional statistical model $M^n = \{p(x;\theta) \mid \theta = (\theta^1, \dots, \theta^n) \in \Theta\}$ is said to be an exponential family, and we say that $\theta = (\theta^1, \dots, \theta^n)$ are its natural parameters. This statistical model M may be viewed as an n–dimensional Riemannian manifold which has natural parameters $(\theta^1, \dots, \theta^n)$ as a local coordinate system. We denote the Fisher metric and the α–connection by g and $\nabla^{(\alpha)}$, respectively. Then the triple $(M, g, \nabla^{(\alpha)})$ is a statistical manifold. Also, the pair $(g, \nabla^{(1)})$ is a Hessian structure ([4]). The multinomial distribution or negative multinomial distribution which are discrete distributions, the multivariate normal distribution, Dirichlet distribution or von Mises-Fisher distribution which are continuous distributions, these distributions are important examples of the exponential family. Especially, the multivariate normal distribution is important on statistics. In [5], L. T. Skovgaard discussed a space of a multivariate normal distribution as a Riemannian manifold. In [7], we treated the statistical submersion with respect to statistical models. Also, we studied geodesics relative to the α–connection such that special spaces of the multivariate normal distribution with a covariance matrix $\mathrm{diag}\,(v_{11}, \dots, v_{nn})$ or $\mathrm{diag}\,(\sigma^2, \dots, \sigma^2)$ in [9].

Also, in [6] we defined a Kähler-like statistical manifold. Let J be an almost complex structure. Then we can define another almost complex structure J^* relative to the Riemannian metric. Moreover, J is parallel with respect to an affine connection ∇ if and only if so is J^* with respect to a conjugate ∇^*. We gave examples of statistical models satisfying these properties in [10]. In [8], we defined an analogy of a Sasakian structure on the statistical manifold.

§2. Statistical manifolds with almost complex structures

Let (M, g) and ∇ be a Riemannian manifold and affine connection, respectively. We define another affine connection ∇^* by

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z) \tag{2.1}$$

for vector fields X, Y and Z on M. An affine connection ∇^* is called conjugate (or dual) of ∇ with respect to g. The triple (M, g, ∇) is called a statistical manifold if both ∇ and ∇^* are torsion-free ([3]). Clearly $(\nabla^*)^* = \nabla$ holds. It is easy to see that $\frac{1}{2}(\nabla + \nabla^*)$ is a metric connection. We denote by R

and R^* the curvature tensors with respect to the affine connection ∇ and its conjugate ∇^*, respectively. Then we find for vector fields X, Y, Z and W

$$g(R(X,Y)Z,W) = -g(Z,R^*(X,Y)W), \tag{2.2}$$

where $R(X,Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z$. Thus R vanishes identically if and only if so is R^*. If the curvature tensor R with respect to the affine connection ∇ satisfies

$$R(X,Y)Z = k\{g(Y,Z)X - g(X,Z)Y\}, \tag{2.3}$$

then the statistical manifold (M, g, ∇) is called a space of constant curvature k.

An almost complex structure on a manifold M is a tensor field J of type (1,1) such that $J^2 = -I$, where I stands for an identity transformation. An almost complex manifold is such a manifold with a fixed almost complex structure. An almost complex manifold is necessarily orientable and must have an even dimension. If J preserves the metric g, that is,

$$g(JX, JY) = g(X,Y) \tag{2.4}$$

for vector fields X and Y on M, then (M, g, J) is an almost Hermitian manifold. Now, we consider the Riemannian manifold (M, g) with an almost complex structure J which has another tensor field J^* of type (1,1) satisfying

$$g(JX, Y) + g(X, J^*Y) = 0. \tag{2.5}$$

Then (M, g, J) is called an almost Hermite-like manifold. We see that $(J^*)^* = J$, $(J^*)^2 = -I$ and

$$g(JX, J^*Y) = g(X,Y). \tag{2.6}$$

If J is parallel with respect to the affine connection ∇, then (M, g, ∇, J) is called a Kähler-like statistical manifold. By virtue of (2.5), we get

$$g((\nabla_Z J)X, Y) + g(X, (\nabla^*_Z J^*)Y) = 0 \tag{2.7}$$

for vector fields X, Y and Z on M. Hence we have ([6])

Lemma A. *(M, g, J) is an almost Hermite-like manifold if and only if so is (M, g, J^*). Moreover, (M, g, ∇, J) is a Kähler-like statistical manifold if and only if so is (M, g, ∇^*, J^*).*

In a Kähler-like statistical manifold, we find

$$R(X,Y)JZ = JR(X,Y)Z. \tag{2.8}$$

If M is of constant curvature k, then we find from (2.3) and (2.8)

$$k\{g(Y,JZ)g(X,W)-g(X,JZ)g(Y,W)\}$$
$$=k\{g(Y,Z)g(JX,W)-g(X,Z)g(JY,W)\}.$$

We assume that $k\neq 0$. Then we obtain $(n-1)g(JX,W)+g(X,JW)-(\operatorname{tr}J)\,g(X,W)=0$, from which $g(JX,W)=g(X,JW)$ if $n>2$. Thus we get $n\,g(JX,W)-(\operatorname{tr}J)\,g(X,W)=0$. Changing X to JX, we find $(\operatorname{tr}J)\,g(JX,W)+n\,g(X,W)=0$. From these two equations, we find $g(X,W)=0$. This is a contradiction. Hence we have

Theorem 2.1. *Let (M^n,g,∇,J) be a Kähler-like statistical manifold. If M $(n\geq 4)$ is of constant curvature, then M is flat.*

§3. α–connection on the statistical model

Let us consider an n–dimensional Riemannian manifold M as a statistical model. For a probability density function $p(x;\theta)$, the parameter x runs through a measure space χ with measure dx so that $\int_\chi p(x;\theta)\,dx=1$ for each θ. The discrete case may be obtained by simply replacing occurrences of the integral $\int_\chi\cdots dx$ with the sum $\sum_{x\in\chi}\cdots$. We put $\ell=\ell(x;\theta)=\log p(x;\theta)$, $\partial_i=\partial/\partial\theta^i$ and we assume that $\partial_1\ell,\ldots,\partial\ell_n$ are linearly independent. We define components of the metric g on M by

$$g_{ij}=E[\,\partial_i\ell\,\partial_j\ell\,], \tag{3.1}$$

where E denotes an expectation relative to $p(x;\theta)$. This metric is independent of the choice of coordinates $(\theta^1,\ldots,\theta^n)$, provided it is finite. It is called a Fisher metric. Since $E[\,\partial_i\ell\,]=0$, it is possible to write g_{ij} as

$$g_{ij}=-E[\,\partial_i\partial_j\ell\,]. \tag{3.2}$$

Also we set functions

$$\Gamma^{(\alpha)}_{ij,k}=E\left[\left(\partial_i\partial_j\ell+\frac{1-\alpha}{2}\partial_i\ell\cdot\partial_j\ell\right)\partial_k\ell\right], \tag{3.3}$$

where α is a real number. We define an α–connection $\nabla^{(\alpha)}$ by

$$g(\nabla^{(\alpha)}_{\partial_i}\partial_j,\partial_k)=\Gamma^{(\alpha)}_{ij,k}. \tag{3.4}$$

Then the α–connection is torsion-free and $\nabla^{(-\alpha)}$ is conjugate of $\nabla^{(\alpha)}$ relative to the Fisher metric. Thus the triple $(M,g,\nabla^{(\alpha)})$ is a statistical manifold. Also, $\nabla^{(0)}$ is the Levi-Civita connection with respect to the Fisher metric. We call α–flat if the curvature tensor with respect to the α–connection vanishes identically.

§4. Exponential families admitting almost complex structures

In this section, we discuss an exponential family admitting almost complex structures which are parallel relative to the exponential connection $\nabla^{(1)}$ or mixture connection $\nabla^{(-1)}$.

The probability density function of the exponential family is given by the equation (1.1). From the normalization condition $\int_\chi p(x;\theta)\,dx = 1$, we find

$$\exp\varphi(\theta) = \int_\chi \exp\left[C(x) + \sum_{s=1}^{n} \theta^s F_s(x) \right] dx. \tag{4.1}$$

We can get $\partial_i\varphi \cdot \exp\varphi = \exp\varphi \cdot E[\,F_i\,]$, which implies that

$$E[\,F_i\,] = \partial_i\varphi. \tag{4.2}$$

Moreover, from $\partial_j(\partial_i\varphi \cdot \exp\varphi) = \exp\varphi \cdot E[\,F_iF_j\,]$ and $\partial_k\{(\partial_i\partial_j\varphi + \partial_i\varphi \cdot \partial_j\varphi)\exp\varphi\} = \exp\varphi \cdot E[\,F_iF_jF_k\,]$, it is easy to see that

$$E[\,F_iF_j\,] = \partial_i\partial_j\varphi + \partial_i\varphi \cdot \partial_j\varphi, \tag{4.3}$$

$$\begin{aligned} E[\,F_iF_jF_k\,] = {} & \partial_k\partial_j\partial_i\varphi + \partial_i\partial_j\varphi \cdot \partial_k\varphi + \partial_j\partial_k\varphi \cdot \partial_i\varphi + \partial_k\partial_i\varphi \cdot \partial_j\varphi \\ & + \partial_i\varphi \cdot \partial_j\varphi \cdot \partial_k\varphi. \end{aligned} \tag{4.4}$$

We set

$$\ell(x;\theta) = \log p(x;\theta) = C(x) + \sum_{s=1}^{n} \theta^s F_s(x) - \varphi(\theta).$$

Owing to (3.2) and $\partial_i\partial_j\ell = -\partial_i\partial_j\varphi$, we have components of the Fisher metric g as follows:

$$g_{ij} = \partial_i\partial_j\varphi. \tag{4.5}$$

Using of (3.3), (4.2), (4.3) and (4.4), we obtain

$$\Gamma^{(\alpha)}_{ij,k} = \frac{1}{2}(1-\alpha)\,\partial_i g_{jk}. \tag{4.6}$$

We put $g^{-1} = (g^{ij})$. Thus we find the α–connection from (3.4)

$$\nabla^{(\alpha)}_{\partial_i}\partial_j = \frac{1}{2}(1-\alpha)\,\partial_s g_{ij} \cdot g^{st}\,\partial_t. \tag{4.7}$$

Then the triple $(M, g, \nabla^{(\alpha)})$ is a statistical manifold. Also, the curvature tensor $R^{(\alpha)}$ relative to the α–connection is rewritten as follows:

$$R^{(\alpha)}(\partial_i, \partial_j)\partial_k = \frac{c(\alpha)}{4}\,(\partial_j g_{ks} \cdot \partial_i g^{st} - \partial_i g_{ks} \cdot \partial_j g^{st})\partial_t, \tag{4.8}$$

where we put $c(\alpha) = (1-\alpha)(1+\alpha)$. S. Amari ([1]) proved that

Theorem B. *The curvature tensor field of the exponential family is given by* (4.8). *Especially, the exponential family is* ± 1*–flat.*

For any real number α, let $J^{(\alpha)}$ be an almost complex structure on M. We seek the condition that the almost complex structure $J^{(\alpha)}$ is parallel with respect to the α–connection. Because of

$$\left(\nabla^{(\alpha)}_{\partial_i} J^{(\alpha)}\right)\partial_j = \left\{\partial_i J_j^{(\alpha)\,t} + \frac{1}{2}(1-\alpha)\left(J_j^{(\alpha)\,r}\partial_i g_{rs}\cdot g^{st} - \partial_i g_{js}\cdot g^{sr} J_r^{(\alpha)\,t}\right)\right\}\partial_t,$$

we find $\nabla^{(\alpha)} J^{(\alpha)} = 0$ is equivalent to following equations

$$\partial_i J_j^{(\alpha)\,k} + \frac{1}{2}(1-\alpha)\left(J_j^{(\alpha)\,r}\partial_i g_{rs}\cdot g^{sk} - \partial_i g_{js}\cdot g^{sr} J_r^{(\alpha)\,k}\right) = 0. \tag{4.9}$$

We consider a system of partial differential equations (4.9) satisfying $J_j^{(\alpha)k}(p) = C_j^{(\alpha)\,k}$ for any $p \in M$ and any constants $C_j^{(\alpha)\,k}$ such that $C_j^{(\alpha)\,r} C_r^{(\alpha)\,k} = -\delta_j{}^k$. We shall show that the system is completely integrable. Letting ∂_h operate on (4.9), we can easily get

$$\begin{aligned}\partial_h\left(\partial_i J_j^{(\alpha)\,k}\right) &+ \frac{1}{2}(1-\alpha)\Big\{J_j^{(\alpha)\,r}(\partial_h\partial_i g_{rs}\cdot g^{sk} + \partial_i g_{rs}\cdot \partial_h g^{sk}) \\ &\qquad - (\partial_h\partial_i g_{js}\cdot g^{sr} + \partial_i g_{js}\cdot\partial_h g^{sr}) J_r^{(\alpha)\,k}\Big\} \\ &+ \frac{1}{4}(1-\alpha)^2\Big\{J_j^{(\alpha)\,r}\partial_h g_{rs}\cdot\partial_i g^{sk} + \partial_i g_{js}\cdot\partial_h g^{sr}\cdot J_r^{(\alpha)\,k} \\ &\qquad + (\partial_h g_{js}\cdot\partial_i g_{tu} + \partial_i g_{js}\cdot\partial_h g_{tu}) g^{sr} J_r^{(\alpha)\,t} g^{uk}\Big\} = 0,\end{aligned}$$

which yields

$$\begin{aligned}&\partial_h\left(\partial_i J_j^{(\alpha)\,k}\right) - \partial_i\left(\partial_h J_j^{(\alpha)\,k}\right) \\ &= -\frac{c(\alpha)}{4}\Big\{J_j^{(\alpha)\,r}(\partial_i g_{rs}\cdot\partial_h g^{sk} - \partial_h g_{rs}\cdot\partial_i g^{sk}) \\ &\qquad - (\partial_i g_{js}\cdot\partial_h g^{sr} - \partial_h g_{js}\cdot\partial_i g^{sr}) J_r^{(\alpha)\,k}\Big\}.\end{aligned}$$

When $\alpha = \pm 1$, we find $\partial_h\left(\partial_i J_j^{(\alpha)\,k}\right) - \partial_i\left(\partial_h J_j^{(\alpha)\,k}\right) = 0$. Thus the system of (4.9) is completely integrable. Also, if $\alpha \neq \pm 1$, then it is easy to see from (4.8) that

$$\partial_h\left(\partial_i J_j^{(\alpha)\,k}\right) - \partial_i\left(\partial_h J_j^{(\alpha)\,k}\right) = -J_j^{(\alpha)\,r} R_{hir}^{(\alpha)\,k} + R_{hij}^{(\alpha)\,r} J_r^{(\alpha)\,k},$$

where $R_{hij}^{(\alpha)\,k}$ are components of the curvature tensor $R^{(\alpha)}$. By virtue of (2.5) and (2.8), we can get $J_j^{(\alpha)\,r} R_{hir}^{(\alpha)\,k} - R_{hij}^{(\alpha)\,r} J_r^{(\alpha)\,k} = 0$, which implies that the system of (4.9) is completely integrable. Hence we have

Theorem 4.1. *The system of partial differential equations* (4.9) *is completely integrable in the exponential family* $(M, g, \nabla^{(\alpha)})$ *for any real number* α.

Especially, if $\alpha = 1$, then we get components of an almost complex structure $J^{(1)}$

$$J_j^{(1)\,k} = P_j{}^k, \tag{4.10}$$

where $P_j{}^k$ are constants satisfying $P_j{}^r P_r{}^k = -\delta_j{}^k$. It is easy to see from (2.5) that $\left(J^{(\alpha)}\right)^* = -g^{-1} J^{(\alpha)} g$. We put

$$J_j^{(-1)\,k} = -P_r{}^s g_{sj} g^{rk}. \tag{4.11}$$

Hence we have

Theorem 4.2. *We find*

(1) $(M, g, J^{(\pm 1)})$ *are almost Hermite-like manifolds*,

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ *are Kähler-like statistical manifolds.*

If $(M, g, \nabla^{(\alpha)})$ is of constant curvature, then the curvature tensor of $\nabla^{(\alpha)}$ can express from (4.7) and Theorem B

$$R_{ijk}^{(\alpha)\,\ell} = c(\alpha)\, A(g_{jk}\delta_i{}^\ell - g_{ik}\delta_j{}^\ell),$$

where A is a constant. Because of Theorem 2.1, we have

Theorem 4.3. *Let* $(M^n, g, \nabla^{(\alpha)})$ $(n \geq 4)$ *be of constant curvature satisfying* $A \neq 0$. *In order for* M *to admit a solution of* (3.9) *such that* $(J^{(\alpha)})^2 = -I$, *it is necessary and sufficient that* $\alpha = \pm 1$.

Remark 4.1. We put $G_{ij}{}^k = \partial_i g_{js} \cdot g^{sk}$. Then we find from (4.9)

$$\partial\, J^{(\alpha)} + \frac{1}{2}(1-\alpha)[\, J^{(\alpha)}, G\,] = 0,$$

where $G = (G_{ij}{}^k)$ and $[\, J^{(\alpha)}, G\,] = J^{(\alpha)} G - G\, J^{(\alpha)}$.

§5. Examples of the exponential family admitting almost complex structures

We consider examples of the discrete or continuous exponential family. We verify exponential families admitting almost complex structures which is parallel relative to $\nabla^{(1)}$ or $\nabla^{(-1)}$.

Example 5.1 (THE MULTINOMIAL DISTRIBUTION). The probability function of a multinomial distribution is given by

$$p(x;\xi) = \frac{N!}{x_1!x_2!\cdots x_{n+1}!}\, p_1^{x_1} p_2^{x_2} \cdots p_{n+1}^{x_{n+1}}, \tag{5.1}$$

where $\xi = (p_1, \dots, p_n)$, $x_k \in \{0, 1, \dots, N\}$ such that $x_1 + x_2 + \cdots + x_{n+1} = N$, and $p_k\,(> 0)$ satisfies $p_1+p_2+\cdots+p_{n+1} = 1$. This probability density function is rewritten as follows:

$$p(x;\xi) = \exp\left(\log N! - \sum_{s=1}^{n+1} \log x_s! + \sum_{s=1}^{n} x_s \log \frac{p_s}{p_{n+1}} + N \log p_{n+1}\right),$$

which implies that the multinomial distribution is an exponential family. We put

$$\begin{aligned}
&C(x) = \log N! - \sum_{s=1}^{n+1} \log x_s!,\\
&F_i(x) = x_i, \qquad \theta^i = \log \frac{p_i}{p_{n+1}} \qquad (i = 1, 2, \dots, n),\\
&\varphi(\theta) = -N \log p_{n+1}
\end{aligned}$$

and $M^n = \{p(x;\theta) \mid \theta = (\theta^1, \dots, \theta^n) \in \mathbb{R}^n\}$. Owing to $p_i = p_{n+1}e^{\theta^i}$ and $p_1 + \cdots + p_{n+1} = 1$, we get $p_{n+1} = \dfrac{1}{\omega(\theta)}$, where we set $\omega(\theta) = 1 + \sum_{s=1}^{n} e^{\theta^s}$, which yields that

$$\varphi(\theta) = N \log \omega(\theta). \tag{5.2}$$

It is clear from (5.2) that

$$\partial_i \varphi = \frac{N\, e^{\theta^i}}{\omega(\theta)}, \tag{5.3}$$

$$\partial_i \partial_j \varphi = N \left\{ \frac{e^{\theta^i}}{\omega(\theta)}\, \delta_{ij} - \frac{e^{\theta^i} e^{\theta^j}}{\omega(\theta)^2} \right\}, \tag{5.4}$$

$$\begin{aligned} \partial_i \partial_j \partial_k \varphi = N \Bigg\{ & \frac{e^{\theta^i}}{\omega(\theta)}\, \delta_{ij}\delta_{ik} - \frac{e^{\theta^i} e^{\theta^k}}{\omega(\theta)^2}\, \delta_{ij} - \frac{e^{\theta^j} e^{\theta^k}}{\omega(\theta)^2}\, \delta_{ik} - \frac{e^{\theta^i} e^{\theta^j}}{\omega(\theta)^2}\, \delta_{jk} \\ & + \frac{2e^{\theta^i} e^{\theta^j} e^{\theta^k}}{\omega(\theta)^3} \Bigg\}, \end{aligned} \tag{5.5}$$

where $\partial_i = \partial/\partial\theta^i$. From (4.5) and (5.4), S. Amari ([1]) calculated the components of the Fisher metric g as follows:

$$g_{ij} = N \left\{ \frac{e^{\theta^i}}{\omega(\theta)}\, \delta_{ij} - \frac{e^{\theta^i} e^{\theta^j}}{\omega(\theta)^2} \right\}. \tag{5.6}$$

Also, components g^{ij} of an inverse matrix of g are given by

$$g^{ij} = \frac{\omega(\theta)}{N\, e^{\theta^i}} (\delta_{ij} + e^{\theta^i}). \tag{5.7}$$

By virtue of (4.6), (5.5) and (5.7), it is easy to see that

$$\Gamma^{(\alpha)\,k}_{ij} = \Gamma^{(\alpha)}_{ij,s}\, g^{sk} = \frac{1}{2}(1-\alpha) \left\{ \delta_{ij}\delta_{ik} - \frac{e^{\theta^j}}{\omega(\theta)}\, \delta_{ik} - \frac{e^{\theta^i}}{\omega(\theta)}\, \delta_{jk} \right\} \tag{5.8}$$

(see [1]). Thus we get

$$\nabla^{(\alpha)}_{\partial_i} \partial_j = \frac{1}{2}(1-\alpha) \left\{ \delta_{ij}\, \partial_i - \frac{e^{\theta^j}}{\omega(\theta)}\, \partial_i - \frac{e^{\theta^i}}{\omega(\theta)}\, \partial_j \right\}. \tag{5.9}$$

The space of a multinomial distribution $(M, g, \nabla^{(\alpha)})$ is a statistical manifold. Moreover we have the curvature tensor relative to the α–connection

$$\begin{aligned} & R^{(\alpha)}(\partial_i, \partial_j)\partial_k \\ & = \frac{c(\alpha)}{4} \left[\left\{ \frac{e^{\theta^j}}{\omega(\theta)}\, \delta_{jk} - \frac{e^{\theta^j} e^{\theta^k}}{\omega(\theta)^2} \right\} \partial_i - \left\{ \frac{e^{\theta^i}}{\omega(\theta)}\, \delta_{ik} - \frac{e^{\theta^i} e^{\theta^k}}{\omega(\theta)^2} \right\} \partial_j \right], \end{aligned}$$

where $c(\alpha) = (1-\alpha)(1+\alpha)$. Hence we have

Theorem 5.1. *The space of a multinomial distribution is of constant curvature* $\frac{c(\alpha)}{4N}$.

We discuss $(M, g, \nabla^{(\alpha)})$ admits an almost complex structure $J^{(\alpha)}$ which is parallel relative to an α–connection $\nabla^{(\alpha)}$. By virtue of Theorems 4.3 and 5.1, we have

Theorem 5.2. *Let* $(M, g, \nabla^{(\alpha)})$ *be a statistical manifold of the multinomial distribution. In order for* M^n $(n \geq 4)$ *to admit almost complex structures* $J^{(\alpha)}$ *which are parallel with respect to* $\nabla^{(\alpha)}$*, it is necessary and sufficient that* $\alpha = \pm 1$.

From (4.10), (4.11), (5.6) and (5.7), we put $J_j^{(1)\,k} = P_j{}^k$ and

$$J_j^{(-1)\,k} = -\frac{e^{\theta^j}}{e^{\theta^k}}\left\{\left(P_k{}^j + e^{\theta^k}\sum_{r=1}^{n} P_r{}^j\right) - \frac{1}{\omega(\theta)}\sum_{s=1}^{n}\left(P_k{}^s + e^{\theta^k}\sum_{r=1}^{n} P_r{}^s\right)e^{\theta^s}\right\},$$

where $P_j{}^k$ are constants satisfying $P_j{}^r P_r{}^k = -\delta_j{}^k$. Hence we have

Theorem 5.3. *In the space of a multinomial distribution, we have*

(1) $(M, g, J^{(\pm 1)})$ *are almost Hermite-like manifolds,*

(2) $(M^n, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ *are Kähler-like statistical manifolds.*

Remark 5.1. If an almost complex structure $J^{(\alpha)}$ on the space of a multinomial distribution is parallel relative to the α–connection, then we find from (4.9), (5.6) and (5.7)

$$\partial_i J_j^{(\alpha)\,k} = \frac{1}{2}(1-\alpha)\left\{(\delta_{ij} - \delta_{ik})J_j^{(\alpha)\,k} - \frac{e^{\theta^j}}{\omega(\theta)}\,J_i^{(\alpha)\,k} + \frac{\delta_{ik}}{\omega(\theta)}\sum_{r=1}^{n} J_j^{(\alpha)\,r} e^{\theta^r}\right\}.$$

When $n = 2$ and $\alpha = 0$, we can get

$$J_1^{(0)\,1} = -J_2^{(0)\,2} = \pm\left(\frac{e^{\theta^1+\theta^2}}{1+e^{\theta^1}+e^{\theta^2}}\right)^{\frac{1}{2}},$$

$$J_1^{(0)\,2} = \pm\frac{1+e^{\theta^2}}{e^{\theta^2}}\left(\frac{e^{\theta^1+\theta^2}}{1+e^{\theta^1}+e^{\theta^2}}\right)^{\frac{1}{2}},$$

$$J_2^{(0)\,1} = \mp\frac{1+e^{\theta^1}}{e^{\theta^1}}\left(\frac{e^{\theta^1+\theta^2}}{1+e^{\theta^1}+e^{\theta^2}}\right)^{\frac{1}{2}}.$$

Therefore $(M^2, g, \nabla^{(0)}, J^{(0)})$ is a Kählerian manifold.

Remark 5.2. Let $\mu = (\mu_1, \dots, \mu_n)$ be a mean vector. From $\mu_i = \frac{N e^{\theta^i}}{\omega(\theta)}$, we get

$$J_j^{(-1)\,k} = -\frac{\mu_j}{\mu_k} \left\{ P_k{}^j + \frac{\mu_k}{N\,p_{n+1}} \sum_{r=1}^{n} P_r{}^j - \frac{1}{N} \sum_{s=1}^{n} \left(P_k{}^s + \frac{\mu_k}{N\,p_{n+1}} \sum_{r=1}^{n} P_r{}^s \right) \mu_s \right\}.$$

Example 5.2 (The Negative Multinomial Distribution). The probability function of a negative multinomial distribution is denoted by

$$p(x;\xi) = \frac{\Gamma(m + x_1 + \cdots + x_n)}{\Gamma(m)\, x_1! x_2! \cdots x_n!}\, p_0^m p_1^{x_1} \cdots p_n^{x_n}, \tag{5.10}$$

where $\xi = (p_1, \dots, p_n)$, $\Gamma(x)$ is the gamma function, m is a positive constant, $x_k \in \{0, 1, 2, \dots\}$ for $k = 1, 2, \dots, n$ and p_k (> 0) satisfies $p_0 + p_1 + \cdots + p_n = 1$. This probability density function is rewritten as follows:

$$p(x;\xi) = \exp\left\{ \log\Gamma(m + x_1 + \cdots + x_n) - \log\Gamma(m) - \sum_{s=1}^{n} \log x_s! + \sum_{s=1}^{n} x_s \log p_s + m\, \log(1 - p_1 - \cdots - p_n) \right\},$$

which means that the negative multinomial distribution is an exponential family. We set

$$C(x) = \log\Gamma(m + x_1 + \cdots + x_n) - \log\Gamma(m) - \sum_{s=1}^{n} \log x_s!,$$
$$F_i(x) = -x_i, \qquad \theta^i = -\log p_i \qquad (i = 1, 2, \dots, n),$$
$$\varphi(\theta) = -m\, \log(1 - p_1 - \cdots - p_n)$$

and $M^n = \{p(x;\theta) \mid \theta = (\theta^1, \dots, \theta^n) \in (\mathbb{R}_+)^n\}$. Because of $p_i = e^{-\theta^i}$ $(i = 1, 2, \dots, n)$, we find

$$\varphi(\theta) = -m\, \log \tau(\theta), \tag{5.11}$$

where we put $\tau(\theta) = 1 - \sum_{s=1}^{n} e^{-\theta^s}$. Therefore we get

$$\partial_i \varphi = -m \, \frac{e^{-\theta^i}}{\tau(\theta)}, \tag{5.12}$$

$$\partial_i \partial_j \varphi = m \left\{ \frac{e^{-\theta^i}}{\tau(\theta)} \, \delta_{ij} + \frac{e^{-\theta^i} e^{-\theta^j}}{\tau(\theta)^2} \right\}, \tag{5.13}$$

$$\begin{aligned} \partial_i \partial_j \partial_k \varphi = -m \left\{ \frac{e^{-\theta^i}}{\tau(\theta)} \, \delta_{ij} \delta_{ik} + \frac{e^{-\theta^i} e^{-\theta^k}}{\tau(\theta)^2} \, \delta_{ij} + \frac{e^{-\theta^j} e^{-\theta^k}}{\tau(\theta)^2} \, \delta_{ik} \right. \\ \left. + \frac{e^{-\theta^i} e^{-\theta^j}}{\tau(\theta)^2} \, \delta_{jk} + \frac{2 e^{-\theta^i} e^{-\theta^j} e^{-\theta^k}}{\tau(\theta)^3} \right\}, \end{aligned} \tag{5.14}$$

where $\partial_i = \partial / \partial \theta^i$. Owing to (4.5) and (5.13), we have components of the Fisher metric g as follows:

$$g_{ij} = m \left\{ \frac{e^{-\theta^i}}{\tau(\theta)} \, \delta_{ij} + \frac{e^{-\theta^i} e^{-\theta^j}}{\tau(\theta)^2} \right\}. \tag{5.15}$$

Also, components g^{ij} of an inverse matrix of g are denoted by

$$g^{ij} = \frac{\tau(\theta)}{m \, e^{-\theta^i}} (\delta_{ij} - e^{-\theta^i}). \tag{5.16}$$

By virtue of (4.6), (5.14) and (5.16), it is clear that following equations hold:

$$\Gamma^{(\alpha)\,k}_{ij} = \Gamma^{(\alpha)}_{ij,s} \, g^{sk} = -\frac{1}{2}(1-\alpha) \left\{ \delta_{ij} \delta_{ik} + \frac{e^{-\theta^j}}{\tau(\theta)} \, \delta_{ik} + \frac{e^{-\theta^i}}{\tau(\theta)} \, \delta_{jk} \right\}. \tag{5.17}$$

Thus we get the following α–connection $\nabla^{(\alpha)}$ for any real number α

$$\nabla^{(\alpha)}_{\partial_i} \partial_j = -\frac{1}{2}(1-\alpha) \left\{ \delta_{ij} \, \partial_i + \frac{e^{-\theta^j}}{\tau(\theta)} \, \partial_i + \frac{e^{-\theta^i}}{\tau(\theta)} \, \partial_j \right\}. \tag{5.18}$$

Therefore the space of a negative multinomial distribution $(M, g, \nabla^{(\alpha)})$ is a statistical manifold. Moreover we find

$$\begin{aligned} & R^{(\alpha)}(\partial_i, \partial_j)\partial_k \\ & = -\frac{c(\alpha)}{4} \left[\left\{ \frac{e^{-\theta^j}}{\tau(\theta)} \, \delta_{jk} + \frac{e^{-\theta^j} e^{-\theta^k}}{\tau(\theta)^2} \right\} \partial_i - \left\{ \frac{e^{-\theta^i}}{\tau(\theta)} \, \delta_{ik} + \frac{e^{-\theta^i} e^{-\theta^k}}{\tau(\theta)^2} \right\} \partial_j \right], \end{aligned}$$

where $c(\alpha) = (1-\alpha)(1+\alpha)$. Hence we have

Theorem 5.4. *The space of a negative multinomial distribution is of constant curvature* $-\frac{c(\alpha)}{4m}$.

Next, we consider $(M, g, \nabla^{(\alpha)})$ admits an almost complex structure $J^{(\alpha)}$ which is parallel relative to an α–connection $\nabla^{(\alpha)}$. From Theorems 4.3 and 5.4, we have

Theorem 5.5. *Let* $(M, g, \nabla^{(\alpha)})$ *be a statistical manifold of the negative multinomial distribution. In order for* M^n $(n \geq 4)$ *to admit almost complex structures* $J^{(\alpha)}$ *which are parallel with respect to* $\nabla^{(\alpha)}$*, it is necessary and sufficient that* $\alpha = \pm 1$.

Taking account of (4.10), (4.11), (5.15) and (5.16), we set $J_j^{(1)\,k} = P_j^{\ k}$ and

$$J_j^{(-1)\,k} = -\frac{e^{-\theta^j}}{e^{-\theta^k}} \left\{ P_k^{\ j} - e^{-\theta^k} \sum_{r=1}^{n} P_r^{\ j} + \frac{1}{\tau(\theta)} \sum_{s=1}^{n} \left(P_k^{\ s} - e^{-\theta^k} \sum_{r=1}^{n} P_r^{\ s} \right) e^{-\theta^s} \right\},$$

where $P_j^{\ k}$ are constants such that $P_j^{\ r} P_r^{\ k} = -\delta_j^{\ k}$. Then we have

Theorem 5.6. *In the space of a negative multinomial distribution, we get*

(1) $(M, g, J^{(\pm 1)})$ *are almost Hermite-like manifolds,*

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ *are Kähler-like statistical manifolds.*

Remark 5.3. If an almost complex structure $J^{(\alpha)}$ on the space of a negative multinomial distribution is parallel relative to the α–connection, then it is easy to see from (4.9), (5.15) and (5.16) that following equations hold

$$\partial_i J_j^{(\alpha)\,k} = \frac{1}{2}(1-\alpha) \left\{ -(\delta_{ij} - \delta_{ik}) J_j^{(\alpha)\,k} - \frac{e^{-\theta^j}}{\tau(\theta)} J_i^{(\alpha)\,k} + \frac{\delta_{ik}}{\tau(\theta)} \sum_{r=1}^{n} J_j^{(\alpha)\,r} e^{-\theta^r} \right\}.$$

When $n = 2$ and $\alpha = 0$, we can get

$$J_1^{(0)\,1} = -J_2^{(0)\,2} = \pm\left(\frac{e^{-\theta^1-\theta^2}}{1-e^{-\theta^1}-e^{-\theta^2}}\right)^{\frac{1}{2}},$$

$$J_1^{(0)\,2} = \mp\frac{1-e^{-\theta^2}}{e^{-\theta^2}}\left(\frac{e^{-\theta^1-\theta^2}}{1-e^{-\theta^1}-e^{-\theta^2}}\right)^{\frac{1}{2}},$$

$$J_2^{(0)\,1} = \pm\frac{1-e^{-\theta^1}}{e^{-\theta^1}}\left(\frac{e^{-\theta^1-\theta^2}}{1-e^{-\theta^1}-e^{-\theta^2}}\right)^{\frac{1}{2}}.$$

Thus $(M^2, g, \nabla^{(0)}, J^{(0)})$ is a Kählerian manifold.

Remark 5.4. By virtue of $\mu_i = \frac{m\,e^{-\theta^i}}{\tau(\theta)}$, we obtain

$$J_j^{(-1)\,k} = -\frac{\mu_j}{\mu_k}\left\{P_k{}^j - \frac{p_0\,\mu_k}{m}\sum_{r=1}^{n}P_r{}^j + \frac{1}{m}\sum_{s=1}^{n}\left(P_k{}^s - \frac{p_0\,\mu_k}{m}\sum_{r=1}^{n}P_r{}^s\right)\mu_s\right\},$$

where μ_i are components of a mean vector.

Example 5.3 (Special Cases of The Multivariate Normal Distribution). The probability density function of a multivariate normal distribution is given by

$$p(x;\xi) = \frac{1}{(\sqrt{2\pi})^n\sqrt{\det\Sigma}}\exp\left[-\frac{1}{2}{}^t(x-\mu)\Sigma^{-1}(x-\mu)\right],$$

where $x = {}^t(x_1,\dots,x_n)$ and $\mu = {}^t(\mu_1,\dots,\mu_n)$ are vectors of order n and μ is called a mean vector, $\Sigma = (\sigma_{ij})$ is a covariance matrix (symmetric positive definite matrix) and $\xi = (\mu_1,\dots,\mu_n,\sigma_{11},\sigma_{12},\dots,\sigma_{1n},\sigma_{22},\dots,\sigma_{2n},\dots,\sigma_{nn}) \in \mathbb{R}^{\frac{1}{2}n(n+3)}$. The multivariate normal distribution is an exponential family. This statistical model may be viewed as a $\frac{1}{2}n(n+3)$–dimensional space which has $(\mu_1,\dots,\mu_n,\sigma_{11},\sigma_{12},\dots,\sigma_{1n},\sigma_{22},\dots,\sigma_{2n},\dots,\sigma_{nn})$ as a local coordinate system. We shall introduce two special spaces of the multivariate normal distribution with the covariance matrix $\Sigma = \mathrm{diag}\,(\sigma_{11},\dots,\sigma_{nn})$ or $\mathrm{diag}\,(\sigma^2,\dots,\sigma^2)$.

At first, we discuss the space of a multivariate normal distribution with the covariance matrix $\mathrm{diag}\,(\sigma_{11},\dots,\sigma_{nn})$. Then the probability density function is denoted by

$$p(x;\xi) = \frac{1}{(\sqrt{2\pi})^n}\prod_{i=1}^{n}\frac{1}{\sqrt{\sigma_{ii}}}\exp\left[-\frac{(x_i-\mu_i)^2}{2\sigma_{ii}}\right],$$

where $\xi = (\mu_1, \ldots, \mu_n, \sigma_{11}, \ldots, \sigma_{nn})$. This statistical model M may be viewed as a $2n$–dimensional space $\mathbb{R}^n \times (\mathbb{R}_+)^n$ which has $(\mu_1, \ldots, \mu_n, \sigma_{11}, \ldots, \sigma_{nn})$ as a local coordinate system. The Fisher metric g and α–connection $\nabla^{(\alpha)}$ were given by [7] and [11]. Furthermore, we proved that the statistical manifold $(M, g, \nabla^{(\alpha)})$ is Einstein. Also, in order for $(M, g, \nabla^{(\alpha)})$ to admit almost complex structures $J^{(\alpha)}$ which is parallel with respect to the α–connection, it is necessary and sufficient that $\alpha = \pm 1$. Therefore $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ are Kähler-like statistical manifolds (see [10]).

Secondly, if the covariance matrix is $\mathrm{diag}\,(\sigma^2, \ldots, \sigma^2)$, then the probability density function of a multivariate normal distribution can be expressed by

$$p(x;\xi) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \prod_{i=1}^{n} \exp\left[-\frac{(x_i - \mu_i)^2}{2\sigma^2}\right],$$

where $\xi = (\mu_1, \ldots, \mu_n, \sigma)$. This statistical model L may be viewed as an $(n+1)$–dimensional upper half-space $\mathbb{R}^n \times \mathbb{R}_+$ which has $(\mu_1, \ldots, \mu_n, \sigma)$ as a local coordinate system. The Fisher metric g and α–connection $\nabla^{(\alpha)}$ were given by [7] and [11]. In [9], the pair (L, g) is a doubly warped product space and the statistical manifold $(L, g, \nabla^{(\alpha)})$ is of a constant curvature $-\frac{c(\alpha)}{2n}$, where $c(\alpha) = (1-\alpha)(1+\alpha)$. Also, for $(L, g, \nabla^{(\alpha)})$, in order to admit almost complex structures $J^{(\alpha)}$ which is parallel relative to the α–connection, it is necessary and sufficient that $\alpha = \pm 1$. Thus $(L, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ are Kähler-like statistical manifolds in [10].

Example 5.4 (The Dirichlet Distribution). The probability density function of the Dirichlet distribution is denoted by

$$p(x;\xi) = \frac{\Gamma(\nu_1 + \cdots + \nu_n)}{\Gamma(\nu_1)\cdots\Gamma(\nu_n)}\, x_1^{\nu_1 - 1} \cdots x_n^{\nu_n - 1}, \tag{5.19}$$

where $\xi = (\nu_1, \ldots, \nu_n)$, $x_k\ (>0)$ satisfies $x_1 + x_2 + \cdots + x_n = 1$ and $\nu_k > 0$ for $k = 1, 2, \ldots, n$. This probability density function is rewritten as follows:

$$p(x;\xi) = \exp\left[-\sum_{s=1}^{n}\{\log x_s - \nu_s \log x_s + \log\Gamma(\nu_s)\} + \log\Gamma(\nu_1 + \cdots + \nu_n)\right],$$

which means that the Dirichlet distribution is an exponential family. We put

$$C(x) = -\sum_{s=1}^{n} \log x_s,$$

$$F_i(x) = \log x_i, \qquad \theta^i = \nu_i \qquad (i = 1, 2, \ldots, n),$$

$$\varphi(\theta) = \sum_{s=1}^{n} \log\Gamma(\nu_s) - \log\Gamma(\nu_1 + \cdots + \nu_n)$$

and $M^n = \{p(x;\theta) \mid \theta = (\theta^1, \dots, \theta^n) \in (\mathbb{R}_+)^n\}$. Thus we get

$$\varphi(\theta) = \sum_{s=1}^{n} \log \Gamma(\theta^s) - \log \Gamma(\theta^1 + \cdots + \theta^n). \tag{5.20}$$

Because of (5.20), it is easy to see that

$$\partial_i \varphi = \psi(\theta^i) - \psi(\theta^1 + \cdots + \theta^n), \tag{5.21}$$

$$\partial_i \partial_j \varphi = \psi'(\theta^i)\, \delta_{ij} - \psi'(\theta^1 + \cdots + \theta^n), \tag{5.22}$$

$$\partial_i \partial_j \partial_k \varphi = \psi''(\theta^i)\, \delta_{ij}\, \delta_{ik} - \psi''(\theta^1 + \cdots + \theta^n), \tag{5.23}$$

where $\partial_i = \partial/\partial\theta^i$ and $\psi(z) = \frac{d}{dz} \log \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ is the digamma function. From (4.5) and (5.22), we have components of the Fisher metric g as follows:

$$g_{ij} = \psi'(\theta^i)\, \delta_{ij} - \psi'(\theta^1 + \cdots + \theta^n). \tag{5.24}$$

Also, components g^{ij} of an inverse matrix of g are given by

$$g^{ij} = \frac{1}{\psi'(\theta^i)} \left\{ \delta_{ij} + \frac{\psi'(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^j)\Psi(\theta^1, \dots, \theta^n)} \right\}, \tag{5.25}$$

where we put

$$\Psi(\theta^1, \dots, \theta^n) = 1 - \psi'(\theta^1 + \cdots + \theta^n) \sum_{s=1}^{n} \frac{1}{\psi'(\theta^s)}. \tag{5.26}$$

By virtue of (4.6), (5.23) and (5.25), we obtain

$$\Gamma^{(\alpha)\,k}_{ij} = \frac{1}{2}(1-\alpha) \left\{ \delta_{ij}\, \delta_{ik} \frac{\psi''(\theta^i)}{\psi'(\theta^i)} + \frac{1}{\psi'(\theta^k)}\, \Phi_{ij}(\theta^1, \dots, \theta^n) \right\}, \tag{5.27}$$

where we set

$$\Phi_{ij}(\theta^1, \dots, \theta^n) = \delta_{ij} \frac{\psi''(\theta^i)\psi'(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^i)\, \Psi(\theta^1, \dots, \theta^n)} - \frac{\psi''(\theta^1 + \cdots + \theta^n)}{\Psi(\theta^1, \dots, \theta^n)}.$$

Thus we get the α–connection $\nabla^{(\alpha)}$

$$\begin{aligned} \nabla^{(\alpha)}_{\partial_i} \partial_j = \frac{1}{2}(1-\alpha) \Bigg\{ & \delta_{ij} \frac{\psi''(\theta^i)}{\psi'(\theta^i)}\, \partial_i \\ & + \Phi_{ij}(\theta^1, \dots, \theta^n) \sum_{s=1}^{n} \frac{1}{\psi'(\theta^s)}\, \partial_s \Bigg\}. \end{aligned} \tag{5.28}$$

Therefore the space of a Dirichlet distribution $(M, g, \nabla^{(\alpha)})$ is a statistical manifold. Moreover we have

$$\begin{aligned}(5.29)\quad & R^{(\alpha)}(\partial_i, \partial_j)\partial_k \\ &= \frac{c(\alpha)}{4}\left\{\Phi_{ik}(\theta^1, \ldots, \theta^n)\,\frac{\psi''(\theta^j)}{\psi'(\theta^j)^2}A_j - \Phi_{jk}(\theta^1, \ldots, \theta^n)\,\frac{\psi''(\theta^i)}{\psi'(\theta^i)^2}A_i\right. \\ &\qquad \left. +(\delta_{jk} - \delta_{ik})\,\frac{\psi''(\theta^k)\psi''(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^k)\,\Psi(\theta^1, \ldots, \theta^n)^2}\sum_{s=1}^{n}\frac{1}{\psi'(\theta^s)}\,\partial_s\right\},\end{aligned}$$

where $c(\alpha) = (1 - \alpha)(1 + \alpha)$ and $A_i = \partial_i + \dfrac{\psi'(\theta^1 + \cdots + \theta^n)}{\Psi(\theta^1, \ldots, \theta^n)}\displaystyle\sum_{s=1}^{n}\frac{1}{\psi'(\theta^s)}\,\partial_s$.

Thus we find

Proposition 5.7. *The curvature tensor field of the space of a Dirichlet distribution is given by* (5.29)*.*

Also, owing to (4.10), (4.11), (5.24) and (5.25), we put $J_j^{(1)\,k} = P_j^{\ k}$ and

$$\begin{aligned}J_j^{(-1)\,k} &= -P_k^{\ j}\,\frac{\psi'(\theta^j)}{\psi'(\theta^k)} + \frac{\psi'(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^k)}\sum_{s=1}^{n}P_k^{\ s} \\ &+\frac{\psi'(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^k)\,\Psi(\theta^1, \ldots, \theta^n)}\sum_{s=1}^{n}\left\{-P_s^{\ j}\,\frac{\psi'(\theta^j)}{\psi'(\theta^s)} + \frac{\psi'(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^s)}\sum_{t=1}^{n}P_s^{\ t}\right\},\end{aligned}$$

where $P_j^{\ k}$ are constants satisfying $P_j^{\ r}P_r^{\ k} = -\delta_j^{\ k}$. Hence we have

Theorem 5.8. *If* $\dim M$ *is even, then we obtain in the space of a Dirichlet distribution*

(1) $(M, g, J^{(\pm 1)})$ *are almost Hermite-like manifolds,*

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ *are Kähler-like statistical manifolds.*

Remark 5.5. If an almost complex structure $J^{(\alpha)}$ on the space of a Dirichlet distribution is parallel relative to the α–connection, then we get

$$\begin{aligned}&\partial_i J_j^{(\alpha)\,k} + \frac{1}{2}(1 - \alpha)\left\{(\delta_{ik} - \delta_{ij})\,\frac{\psi''(\theta^i)}{\psi'(\theta^i)}\,J_j^{(\alpha)\,k}\right. \\ &\left.+\frac{1}{\psi'(\theta^k)}\sum_{s=1}^{n}J_j^{(\alpha)\,s}\,\Phi_{is}(\theta^1, \ldots, \theta^n) - \Phi_{ij}(\theta^1, \ldots, \theta^n)\sum_{s=1}^{n}\frac{J_s^{(\alpha)\,k}}{\psi'(\theta^s)}\right\} = 0,\end{aligned}$$

where we have used (4.9), (5.24) and (5.25).

Example 5.5 (The von Mises-Fisher Distribution). Let S^{n-1} be an unite sphere of $\mathbb{R}^n$. For $\mu = (\mu_1, \ldots, \mu_n) \in S^{n-1}$ and $x = (x_1, \ldots, x_n) \in S^{n-1}$, we put $\langle \mu, x \rangle = \displaystyle\sum_{s=1}^{n} \mu_s x_s$. The probability density function of the von Mises-Fisher distribution is denoted by

$$p(x;\xi) = C_n(\kappa) \exp\Big(\kappa \langle \mu, x \rangle\Big), \tag{5.30}$$

where $\kappa > 0$ and $\xi = (\mu_1, \ldots, \mu_n, \kappa)$. This is a probability distribution on the sphere S^{n-1} and

$$C_n(\kappa) = \frac{\kappa^p}{(2\pi)^{p+1} I_p(\kappa)}, \tag{5.31}$$

where $p = (n-2)/2$ and $I_p(\kappa)$ is the modified Bessel function of the first kind. This probability density function is rewritten as follows:

$$p(x;\xi) = \exp\{\kappa \langle \mu, x \rangle + \log C_n(\kappa)\},$$

which means that the von Mises-Fisher distribution is an exponential family. We get $C(x) = 0$,

$$\begin{aligned} &F_i(x) = x_i, \qquad \theta^i = \kappa\, \mu_i \qquad (i = 1, 2, \ldots, n), \\ &\varphi(\theta) = -\log C_n(\kappa) \end{aligned}$$

and $M^n = \{p(x;\theta) \mid \theta = (\theta^1, \ldots, \theta^n) \in \mathbb{R}^n\}$. This manifold M^n is a product of the sphere S^{n-1} and the half line $\mathbb{R}_+$. From $\mu \in S^{n-1}$, we find $\kappa = \langle \theta, \theta \rangle^{\frac{1}{2}}$. Thus we can get

$$\varphi(\theta) = \log I_p(\kappa) - p \log \kappa + (p+1) \log 2\pi. \tag{5.32}$$

Because of (5.32), it is easy to see from following equations with respect to the modified Bessel function of the first kind

$$\begin{aligned} I_p'(\kappa) &= \frac{1}{2}\Big\{I_{p-1}(\kappa) + I_{p+1}(\kappa)\Big\}, \\ \frac{2p}{\kappa}\, I_p(\kappa) &= I_{p-1}(\kappa) - I_{p+1}(\kappa) \end{aligned}$$

that

$$\partial_i \varphi = F_p(\kappa) \theta^i, \tag{5.33}$$

$$\partial_i \partial_j \varphi = F_p(\kappa)\, \delta_{ij} + \frac{F_p'(\kappa)}{\kappa}\, \theta^i \theta^j, \tag{5.34}$$

$$\partial_i \partial_j \partial_k \varphi = \frac{F_p'(\kappa)}{\kappa}(\delta_{ij}\theta^k + \delta_{jk}\theta^i + \delta_{ki}\theta^j) + \frac{1}{\kappa}\left(\frac{F_p'(\kappa)}{\kappa}\right)' \theta^i \theta^j \theta^k, \tag{5.35}$$

where $\partial_i = \partial/\partial\theta^i$ and

$$F_p(\kappa) = \frac{I_{p+1}(\kappa)}{\kappa\, I_p(\kappa)}.$$

From (4.5) and (5.34), we have components of the Fisher metric g as follows:

$$g_{ij} = F_p(\kappa)\,\delta_{ij} + \frac{F_p'(\kappa)}{\kappa}\,\theta^i\theta^j. \tag{5.36}$$

Also, components g^{ij} of an inverse matrix of g are given by

$$g^{ij} = \frac{1}{F_p(\kappa)}\,\delta_{ij} - \frac{F_p'(\kappa)}{\kappa\, F_p(\kappa)\,\{\kappa\, F_p(\kappa)\}'}\,\theta^i\theta^j. \tag{5.37}$$

By virtue of (4.6), (5.36) and (5.37), we obtain

$$\begin{aligned}\Gamma^{(\alpha)\,k}_{ij} = \frac{1-\alpha}{2F_p(\kappa)} \Bigg\{ & F_p'(\kappa)\,(\delta_{jk}\,\theta^i + \delta_{ik}\,\theta^j) \\ & + F_p'(\kappa)\left(1 - \frac{\kappa}{\{\kappa\, F_p(\kappa)\}'}\right)\delta_{ij}\theta^k \\ & + \frac{F_p(\kappa)\left(\frac{F_p'(\kappa)}{\kappa}\right)' - 2F_p'(\kappa)^2}{\kappa\,\{\kappa\, F_p(\kappa)\}'}\,\theta^i\theta^j\theta^k \Bigg\}.\end{aligned} \tag{5.38}$$

Therefore we obtain

$$\begin{aligned}\nabla^{(\alpha)}_{\partial_i}\partial_j = \frac{1-\alpha}{2F_p(\kappa)} \Bigg\{ & F_p'(\kappa)\,(\theta^i\,\partial_j + \theta^j\,\partial_i) \\ & + F_p'(\kappa)\left(1 - \frac{\kappa}{\{\kappa\, F_p(\kappa)\}'}\right)\delta_{ij}\sum_{s=1}^{n}\theta^s\partial_s \\ & + \frac{F_p(\kappa)\left(\frac{F_p'(\kappa)}{\kappa}\right)' - 2F_p'(\kappa)^2}{\kappa\,\{\kappa\, F_p(\kappa)\}'}\,\theta^i\theta^j\sum_{s=1}^{n}\theta^s\partial_s \Bigg\}.\end{aligned} \tag{5.39}$$

Thus the space of a von Mises-Fisher distribution $(M, g, \nabla^{(\alpha)})$ is a statistical manifold. Moreover we have

$$\begin{aligned} & R^{(\alpha)}(\partial_i, \partial_j)\partial_k \\ & = \frac{c(\alpha)}{4}\Bigg[\frac{F_p'(\kappa)}{\kappa^2 F_p(\kappa)}\left\{1 - \frac{F_p'(\kappa)}{F_p(\kappa)}\right\}\theta^k(\theta^i\partial_j - \theta^j\partial_i) \\ & \quad + \frac{2\kappa\, F_p'(\kappa)^2}{\{\kappa^2 F_p(\kappa)^2\}'}(\delta_{ik}\partial_j - \delta_{jk}\partial_i) \\ & \quad - H_p(\kappa)\,(\delta_{jk}\theta^i - \delta_{ik}\theta^j)\sum_{s=1}^{n}\theta^s\partial_s\Bigg],\end{aligned} \tag{5.40}$$

where we put

$$H_p(\kappa) = \left\{ \frac{F_p'(\kappa)}{\kappa\, F_p(\kappa)} \right\}^2 + F_p'(\kappa) \left\{ \frac{2F_p'(\kappa)}{\{\kappa^2 F_p(\kappa)^2\}'} \right\}'.$$

Thus we find

Proposition 5.9. *The curvature tensor field of the space of a von Mises-Fisher distribution is given by* (4.39).

Also, by virtue of (4.10), (4.11), (5.36) and (5.37), we put $J_j^{(1)\,k} = P_j^{\ k}$ and

$$\begin{aligned} J_j^{(-1)\,k} &= -P_k^{\ j} - \frac{F_p'(\kappa)}{\kappa\, F_p(\kappa)}\, \theta^j \sum_{s=1}^n P_k^{\ s} \theta^s \\ &\quad + \frac{F_p'(\kappa)}{\kappa \{\kappa\, F_p(\kappa)\}'}\, \theta^k \sum_{r=1}^n \left\{ P_r^{\ j} + \frac{F_p'(\kappa)}{\kappa\, F_p(\kappa)}\, \theta^j \sum_{s=1}^n P_r^{\ s} \theta^s \right\} \theta^r, \end{aligned}$$

where $P_j^{\ k}$ are constants satisfying $P_j^{\ r} P_r^{\ k} = -\delta_j^{\ k}$. Hence we have

Theorem 5.10. *If* $\dim M$ *is even, then we obtain in the space of a von Mises-Fisher distribution*

(1) $(M, g, J^{(\pm 1)})$ *are almost Hermite-like manifolds,*

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ *are Kähler-like statistical manifolds.*

Remark 5.6. If an almost complex structure $J^{(\alpha)}$ on the space of a von Mises-Fisher distribution is parallel with respect to $\nabla^{(\alpha)}$, then we can get

$$\begin{aligned} \partial_i J_j^{(\alpha)\,k} + \frac{1-\alpha}{2\kappa\, F_p(\kappa)} &\left[F_p'(\kappa) \left\{ 1 - \frac{\kappa\, F_p'(\kappa)}{\{\kappa\, F_p(\kappa)\}'} \right\} \left(\theta^k J_j^{(\alpha)\,i} - \delta_{ij} \sum_{s=1}^n \theta^s J_s^{(\alpha)\,k} \right) \right. \\ &\quad - F_p'(\kappa) \left(\theta^j J_i^{(\alpha)\,k} - \delta_{ik} \sum_{s=1}^n J_j^{(\alpha)\,s} \theta^s \right) \\ &\quad \left. + K_p(\kappa)\, \theta^i \sum_{s=1}^n \left(\theta^k \theta^s J_j^{(\alpha)\,s} - \theta^j \theta^s J_s^{(\alpha)\,k} \right) \right] = 0, \end{aligned}$$

where we put

$$K_p(\kappa) = \left(\frac{F_p'(\kappa)}{\kappa} \right)' - \frac{2F_p'(\kappa)^2}{\kappa\, \{\kappa\, F_p(\kappa)\}'} - \frac{\kappa\, F_p'(\kappa)}{\{\kappa\, F_p(\kappa)\}'} \left(\frac{F_p'(\kappa)}{\kappa} \right)'.$$

References

[1] S. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics, **28** Springer-Verlag, 1985.

[2] S. Amari and H. Nagaoka, *Methods of Information Geometry*, AMS & Oxford University Press, 2000.

[3] M. Noguchi, *Geometry of statistical manifolds*, Differential Geom. Appl. **2** (1992), 197–222.

[4] H. Shima, *The geometry of Hessian structures*, World Scientific, 2007.

[5] L. T. Skovgaard, *A Riemannian geometry of the multivariate normal model*, Scand. J. Statist. **11** (1984), 211–223.

[6] K. Takano, *Statistical manifolds with almost complex structures and its statistical submersions*, Tensor, N. S. **65** (2004), 128–142.

[7] K. Takano, *Examples of the statistical submersion on the statistical model*, Tensor, N. S. **65** (2004), 170–178.

[8] K. Takano, *Statistical manifolds with almost contact structures and its statistical submersions*, J. Geom. **85** (2006), 171–187.

[9] K. Takano, *Geodesics on statistical models of the multivariate normal distribution*, Tensor, N. S. **67** (2006), 162–169.

[10] K. Takano, *Examples of statistical manifolds with almost complex structures*, Tensor, N. S. **69** (2008), 58–66.

[11] M. Yamada, R. Ivanova, Y. Sasahara and T. Kawaguchi, *On some properties of the statistical parameter space of the multivariate normal distribution*, Tensor, N. S. **59** (1998), 77–82.

[12] K. Yano and M. Kon, *Structures on Manifolds*, World Scientific, 1984.

Kazuhiko Takano
Department of Mathematics
School of General Education
Shinshu University
Matsumoto 390-8621, Japan
E-mail: `ktakano@shinshu-u.ac.jp`