大阪市立大学数学研究所ミニスクール
情報幾何への入門と応用,
平成 19 年 12 月 21 日

指数型分布族の概複素構造について

信州大学全学教育機構
高野　嘉寿彦

1　はじめに

情報幾何学における統計的モデルは，期待値から構成される Fisher 計量 g，α–接続とよばれる捩れのないアファイン接続 $\nabla^{(\alpha)}$ を許容する。ここで，α は実数である。Fisher 計量に関して α–接続と $(-\alpha)$–接続は互いに共役であり，0–接続は Levi–Civita 接続である。特に，$\nabla^{(1)}$ 及び $\nabla^{(-1)}$ はそれぞれ指数型接続及び混合型接続，または簡単に e–接続及び m–接続と呼ばれ $\nabla^{(e)}$ 及び $\nabla^{(m)}$ で表す。これらは情報幾何学において重要な概念である ([1], [2])。

確率密度関数が集合 χ 上の関数 $C, F_1, \ldots, F_n$ と $\mathbf{R}^n$ の部分集合 Θ 上の関数 φ を用いて

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

と表せるとき，n 次元統計的モデル $M = \{p(x;\theta) \mid \theta = (\theta^1, \ldots, \theta^n) \in \Theta\}$ を指数型分布族といい，$\theta = (\theta^1, \ldots, \theta^n)$ をその自然パラメータという。統計的モデル M は局所座標系として自然パラメータ $(\theta^1, \ldots, \theta^n)$ をもつ n 次元 Riemann 多様体と考えられる。$(M^n, g, \nabla^{(\alpha)})$ は統計多様体になる。離散型分布である多項分布や負の多項分布，連続型分布である多次元正規分布や Dirichlet 分布は指数型分布族の重要な例である。今回，これらの統計多様体の概複素構造について述べる。

2　概複素構造をもつ統計多様体

(M, g) を Riemann 多様体とし，アファイン接続を ∇ で表す。M 上のベクトル場 X, Y, Z に対して

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

によりもう一つのアファイン接続 ∇^* を定義する。この ∇^* を g に関する ∇ の共役接続 (双対接続) という。$(\nabla^*)^* = \nabla$ を満たす。∇ と ∇^* の捩れがないとき，組 (M, g, ∇) を統計多様体という ([3])。また，組 (M, g, ∇^*) も統計多様体になる。アファイン接続 ∇ 及び ∇^* に関する曲率テンソルをそれぞれ R 及び R^* で表す。このとき，M 上のベクトル場 X, Y, Z, W に対して

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

が成り立つ。ここで，$R(X,Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z$ である。これより R が恒等的に零であることと R^* が恒等的に零であることは同値である。また，アファイン接続 ∇ に関する曲率テンソル R が

$$R(X,Y)Z = c\{\, g(Y,Z)X - g(X,Z)Y \,\} \qquad (c \text{ は定数}) \tag{2.3}$$

を満たすとき統計多様体 (M,g,∇) をアファイン接続 ∇ に関する定曲率空間という。(M,g,∇) が ∇ に関して定曲率空間であることと (M,g,∇^*) が ∇^* に関して定曲率空間であることは同値である。

多様体 M 上の概複素構造とは，$J^2=-I$ となる (1,1) 型のテンソル場 J である。ここで，I は恒等変換を表す。固定されたそのような概複素構造をもつ多様体を概複素多様体という。概複素多様体は向き付け可能であり偶数次元である。M が Riemann 多様体である概複素多様体 (M,J) を考える。M 上のベクトル場 X, Y に対して

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

を満たすとき，(M,g,J) を概 Hermite 多様体という。次に

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

によりもう一つの (1,1) 型のテンソル場 J^* を定義する。このような Riemann 多様体 (M,g) を考える。このとき (M,g,J) を概 Hermite-like 多様体とよぶ。$(J^*)^*=J$，$(J^*)^2=-I$ 及び

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

が成り立つ。このような統計多様体において，J がアファイン接続 ∇ に関して平行であるとき，(M,g,∇,J) を Kähler-like 統計多様体と定める。(2.5) よりベクトル場 X, Y, Z に対して

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

が得られる。これより ([5])

補題 2.1. 次が成り立つ。

(1) (M,g,J) が概 Hermite-like 多様体であること (M,g,J^*) が概 Hermite-like 多様体であることは同値である。

(2) (M,g,∇,J) が Kähler-like 統計多様体であることと (M,g,∇^*,J^*) が Kähler-like 統計多様体であることは同値である。

3　統計的モデルについて

統計的モデルとしての n 次元 Riemann 多様体 M を考える。確率密度関数 $p(x;\theta)$ に対して $\int_\chi p(x;\theta)\,dx=1$ である。ここで確率分布が離散型のときは積分は和 $\sum_{x\in\chi}\cdots$ になる。$\ell=\ell(x;\theta)=\log p(x;\theta)$ 及び $\partial_i=\partial/\partial\theta^i$ とおき，$\partial_1\ell,\dots,\partial_n\ell$ は一次独立とし，微分と積分の順序交換が可能と仮定する。$p(x;\theta)$ に関する期待値 E を用いて M 上の計量 g の成分を

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

で定義する。この計量は局所座標系 $(\theta^1,\dots,\theta^n)$ の取り方に依存しない。この g を Fisher 計量という。$E[\,\partial_i\ell\,]=0$ より

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

と表すこともできる。

また，実数 α に対して関数を

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

と定める。これにより α–接続 $\nabla^{(\alpha)}$ を

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

で定義する。このとき α–接続は捩れがなく，Fisher 計量に関して $\nabla^{(\alpha)}$ と $\nabla^{(-\alpha)}$ は互いに共役である。これより $(M, g, \nabla^{(\alpha)})$ は統計多様体である。また，$\nabla^{(0)}$ は Fisher 計量に関する Levi-Civita 接続である。α–接続に関する曲率テンソルが恒等的に零のとき α–平坦という。

4　指数型分布族の概複素構造

指数型分布族の確率密度関数は式 (1.1) で表せる。連続型とすると正規化条件 $\int_\chi p(x;\theta)\,dx = 1$ より

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

これを θ^i で微分すると $\partial_i\varphi\cdot\exp\varphi = \exp\varphi\cdot E[\,F_i\,]$ が得られるから

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

が分かる。さらに微分すると $\partial_j(\partial_i\varphi\cdot\exp\varphi) = \exp\varphi\cdot E[\,F_iF_j\,]$ 及び $\partial_k\{(\partial_i\partial_j\varphi + \partial_i\varphi\cdot\partial_j\varphi)\exp\varphi\} = \exp\varphi\cdot E[\,F_iF_jF_k\,]$ より

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

$$E[\,F_iF_jF_k\,] = \partial_k\partial_j\partial_i\varphi + \partial_i\partial_j\varphi\cdot\partial_k\varphi + \partial_j\partial_k\varphi\cdot\partial_i\varphi + \partial_k\partial_i\varphi\cdot\partial_j\varphi + \partial_i\varphi\cdot\partial_j\varphi\cdot\partial_k\varphi \tag{4.4}$$

が得られる。また式 (1.1) において対数をとり

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

とおく。式 (3.2) と $\partial_i\partial_j\ell = -\partial_i\partial_j\varphi$ から Fisher 計量 g の成分は次のようになる：

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

また $g^{-1} = (g^{ij})$ とする。式 (3.3), (4.2), (4.3), (4.4) より

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

を得る。これより式 (3.4) から α–接続

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

が得られ，$(M, g, \nabla^{(\alpha)})$ は統計多様体になる。また α–接続に関する曲率テンソル $R^{(\alpha)}$ は次のように表せる：

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

ここで $c(\alpha) = (1-\alpha)(1+\alpha)$ とおいた。これより

命題 4.1． 指数型分布族の曲率テンソルは (4.8) で与えられる。特に ± 1–平坦である。

実数 α に対して $J^{(\alpha)}$ を M 上の概複素構造とする。そのとき (2.5) を満たすもう一つの概複素構造は $J^{(-\alpha)}$ である。α–接続に関して概複素構造 $J^{(\alpha)}$ が平行になる条件を求める。

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

より，$\nabla^{(\alpha)} J^{(\alpha)} = 0$ は次と同値である：

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

これより $\alpha = 1$ ならば $J^{(1)}$ の成分は

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

となる。ここで $P_j{}^k$ は $P_j{}^r P_r{}^i = -\delta_j{}^i$ を満たす定数である。$g(J^{(1)}\partial_i, \partial_j) + g(\partial_i, J^{(-1)}\partial_j) = 0$, 即ち，$J_i^{(1)\,r} g_{rj} + J_j^{(-1)\,r} g_{ir} = 0$ から $J^{(-1)}$ の成分

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

が得られる。これより

定理 4.2． 指数型分布族において

(1) $(M, g, J^{(\pm 1)})$ はそれぞれ概 Hermite-like 多様体である。

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ はそれぞれ Kähler-like 統計多様体である。

注意 4.1． $G_{ij}{}^k = \partial_i g_{js}\cdot g^{sk}$ とおくと，式 (4.9) は

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

と表せる。ここで，$G = (G_{ij}{}^k)$ および $[\, J^{(\alpha)}, G\,] = J^{(\alpha)}G - G\, J^{(\alpha)}$ である。

5　概複素構造をもつ指数型分布族の例

具体的な離散型や連続型の指数型分布族を考え，1–接続や (-1)–接続に関して平行である概複素構造を与える。

例 5.1 (多項分布)． 多項分布の確率密度関数は

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

で与えられる。ここで，$\xi = (p_1, \ldots, p_n)$ であり，$x_k \in \{0, 1, \ldots, N\}$ は $x_1 + x_2 + \cdots + x_{n+1} = N$ を満たし，$p_k\,(>0)$ は $p_1 + p_2 + \cdots + p_{n+1} = 1$ を満たす。これは次のように表すことができる：

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

これより多項分布は指数型である。

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

$$F_i(x) = x_i, \quad \theta^i = \log \frac{p_i}{p_{n+1}} \quad (i = 1, 2, \dots, n),$$

$$\varphi(\theta) = -N \log p_{n+1}$$

とおき，$M^n = \{p(x;\theta) \mid \theta = (\theta^1, \dots, \theta^n) \in \mathbf{R}^n\}$ とする。$p_i = p_{n+1} e^{\theta^i}$ 及び $p_1 + \dots + p_{n+1} = 1$ より $p_{n+1} = \dfrac{1}{\omega(\theta)}$. ここで $\omega(\theta) = 1 + \sum_{s=1}^{n} e^{\theta^s}$ とおいた。これより

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

となる。これを微分すると

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

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

$$\partial_i \partial_j \partial_k \varphi = N \left\{ \frac{e^{\theta^i}}{\omega(\theta)} \delta_{ij}\delta_{ik} - \frac{e^{\theta^i} e^{\theta^k}}{\omega(\theta)^2} \delta_{ij} - \frac{e^{\theta^j} e^{\theta^k}}{\omega(\theta)^2} \delta_{ik} - \frac{e^{\theta^i} e^{\theta^j}}{\omega(\theta)^2} \delta_{jk} + \frac{2 e^{\theta^i} e^{\theta^j} e^{\theta^k}}{\omega(\theta)^3} \right\} \tag{5.5}$$

が得られる。ここで $\partial_i = \partial / \partial \theta^i$ である。式 (4.5) 及び (5.4) より多項分布のなす空間の Fisher 計量は

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

となる。また g の逆行列の成分 g^{ij} は次で与えられる：

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

式 (4.6), (5.5), (5.7) から

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

が得られる。これより α–接続 $\nabla^{(\alpha)}$ は

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

となる。多項分布のなす空間 $(M, g, \nabla^{(\alpha)})$ は統計多様体であり，α–接続に関する曲率テンソルは

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

となる。ただし $c(\alpha) = (1-\alpha)(1+\alpha)$ である。これより

定理 5.1． 多項分布のなす空間は定曲率 $\frac{c(\alpha)}{4N}$ の定曲率空間である。

次に (4.10), (4.11), (5.6), (5.7) から 1–接続及び (-1)–接続に関して平行となる概複素構造 $J^{(1)}$ 及び $J^{(-1)}$ の成分はそれぞれ $J_j^{(1)\,k} = P_j^{\ k}$ 及び

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

で与えられる。これより

定理 5.2. 多項分布のなす空間において

(1) $(M, g, J^{(\pm 1)})$ はそれぞれ概 Hermite-like 多様体である。

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ はそれぞれ Kähler-like 統計多様体である。

例 5.2 (**負の多項分布**). 負の多項分布の確率密度関数は

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

で表される。ここで，$\xi = (p_1, \ldots, p_n)$，$\Gamma(x)$ はガンマ関数，m は正定数であり，$k = 1, 2, \ldots, n$ に対して $x_k \in \{0, 1, 2, \ldots\}$，$p_k\,(> 0)$ は $p_0 + p_1 + \cdots + p_n = 1$ を満たす。この確率分布は

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

と表すことができるから負の多項分布は指数型分布族である。

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

とおき，$M^n = \{p(x;\theta) \mid \theta = (\theta^1, \ldots, \theta^n) \in (\mathbf{R}_+)^n\}$ とする。$p_i = e^{-\theta^i}$ $(i = 1, 2, \ldots, n)$ より

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

ここで $\tau(\theta) = 1 - \sum_{s=1}^{n} e^{-\theta^s}$ とおいた。これを微分すると

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

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

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

が得られる。式 (4.5) と (5.13) から Fisher 計量は次のようになる：

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

また Fisher 計量 g の逆行列の成分は

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

となることが分かる。式 (4.6), (5.14), (5.16) から次式が成り立つ:

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

これより，実数 α に対して α–接続 $\nabla^{(\alpha)}$ は

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

となるから負の多項分布のなす空間は $(M, g, \nabla^{(\alpha)})$ は統計多様体である。さらに α–接続に関する曲率テンソルは

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

となる。ここで $c(\alpha) = (1-\alpha)(1+\alpha)$ とおいた。これより

定理 5.3． 負の多項分布のなす空間は定曲率 $-\frac{c(\alpha)}{4m}$ の定曲率空間である。

次に，式 (4.10), (4.11), (5.15), (5.16) から 1–接続及び (-1)–接続に関して平行となる概複素構造 $J^{(1)}$ 及び $J^{(-1)}$ の成分はそれぞれ $J_j^{(1)\,k} = P_j{}^k$ 及び

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

と表すことができる。これより

定理 5.4． 負の多項分布のなす空間において

(1) $(M, g, J^{(\pm 1)})$ はそれぞれ概 Hermite-like 多様体である。

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ はそれぞれ Kähler-like 統計多様体である。

例 5.3 (多次元正規分布の特別な場合)．多次元正規分布の確率密度関数は

$$p(x;\xi) = \frac{1}{(\sqrt{2\pi})^n\sqrt{v}}\exp\left[-\frac{1}{2}{}^t(x-m)V^{-1}(x-m)\right]$$

で与えられる。ここで $x = {}^t(x_1, \ldots, x_n)$, $m = {}^t(m_1, \ldots, m_n)$ はベクトルであり，m は平均ベクトルと呼ばれる。$V = (v_{ij})$ は共分散行列と呼ばれる対称な正定値行列であり，$v = \det V$ 及び $\xi = (m_1, \ldots, m_n, v_{11}, v_{12}, \ldots, v_{1n}, v_{22}, \ldots, v_{2n}, \ldots, v_{nn}) \in \mathbf{R}^{\frac{1}{2}n(n+3)}$ である。多次元正規分布は指数型であり，ξ を局所座標系として $\mathbf{R}^{\frac{1}{2}n(n+3)}$ の $\frac{1}{2}n(n+3)$ 次元部分空間である。共分散行列 V が $\mathrm{diag}\,(v_{11}, \ldots, v_{nn})$ や $\mathrm{diag}\,(\sigma^2, \ldots, \sigma^2)$ であるような特別な多次元正規分布を考える。

最初に，共分散行列が $\mathrm{diag}\,(v_{11}, \ldots, v_{nn})$ である多次元正規分布の確率密度関数は

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

のように表せる。ここで $\xi = (m_1, \ldots, m_n, v_{11}, \ldots, v_{nn})$ である。この統計的モデル M は ξ を局所座標系として $2n$ 次元空間 $\mathbf{R}^n \times (\mathbf{R}_+)^n$ である。Fisher 計量 g 及び α–接続 $\nabla^{(\alpha)}$ は [6] や [11] で与えられた。さらに統計多様体 $(M, g, \nabla^{(\alpha)})$ は Einstein であることが示された ([6])。また，$(M, g, \nabla^{(\alpha)})$ が α–接続に関して平行となる概複素構造 $J^{(\alpha)}$ を許容するとき $\alpha = \pm 1$ であることが示され，これより $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ はそれぞれ Kähler-like 統計多様体になる ([9])。

次に，共分散行列が $\mathrm{diag}\,(\sigma^2, \ldots, \sigma^2)$ のとき，確率密度関数は次のように表すことができる：

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

ここで $\xi = (m_1, \ldots, m_n, \sigma)$ である。この統計的モデル L は局所座標系 $(m_1, \ldots, m_n, \sigma)$ をもつ $(n+1)$ 次元上半空間 $\mathbf{R}^n \times \mathbf{R}_+$ になる。Fisher 計量 g や α–接続 $\nabla^{(\alpha)}$ は [8] で与えられ，統計多様体 $(L, g, \nabla^{(\alpha)})$ は定曲率 $-\frac{c(\alpha)}{2n}$ の定曲率空間であることを示した。ただし $c(\alpha) = (1-\alpha)(1+\alpha)$ である。また，$(L, g, \nabla^{(\alpha)})$ が α–接続に関して平行となる概複素構造 $J^{(\alpha)}$ を許容するとき $\alpha = \pm 1$ である。これより $(L, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ はそれぞれ Kähler-like 統計多様体になる ([9])。

例 5.4 (**Dirichlet 分布**). Dirichlet 分布の確率密度関数は

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

である。ここで，$\xi = (\nu_1, \ldots, \nu_n)$，正数 x_k は $x_1 + x_2 + \cdots + x_n = 1$ を満たし，$k = 1, 2, \ldots, n$ に対して $\nu_k > 0$ である。$n = 2$ のとき，ベータ分布である。この確率密度関数は次のように書きなおせる：

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

これより Dirichlet 分布は指数型である。

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

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

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

とおき，$M^n = \{p(x;\theta) \mid \theta = (\theta^1, \ldots, \theta^n) \in (\mathbf{R}_+)^n\}$ とする。このとき

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

であり，これを微分すると

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

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

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

が得られる。$\partial_i = \partial/\partial\theta^i$ であり，$\psi(\theta) = \frac{d}{d\theta}\log\Gamma(\theta) = \frac{\Gamma'(\theta)}{\Gamma(\theta)}$ は digamma 関数を表す。式 (4.5) と (5.22) から Fisher 計量 g の成分は

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

となり，g の逆行列の成分 g^{ij} は

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

で与えられる。ここで

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

とおいた。式 (4.6), (5.23), (5.25) から

$$\begin{aligned}\Gamma_{ij}^{(\alpha)\,k} = \frac{1}{2}(1-\alpha)\,\delta_{ij}\,\frac{\psi''(\theta^i)}{\psi'(\theta^i)}\left\{\delta_{ik} + \frac{\psi'(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^k)\Psi(\theta^1, \ldots, \theta^n)}\right\} \\ -\frac{1}{2}(1-\alpha)\,\frac{\psi''(\theta^1 + \cdots + \theta^n)}{\psi'(\theta^k)\Psi(\theta^1, \ldots, \theta^n)}\end{aligned} \tag{5.27}$$

だから α–接続

$$\nabla^{(\alpha)}_{\partial_i}\partial_j = \frac{1}{2}(1-\alpha)\left\{\delta_{ij}\,\frac{\psi''(\theta^i)}{\psi'(\theta^i)}\,\partial_i + \Phi_{ij}(\theta^1, \ldots, \theta^n)\,\sum_{s=1}^{n}\frac{1}{\psi'(\theta^s)}\,\partial_s\right\} \tag{5.28}$$

が得られる。ここで

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

である。これより Dirichlet 分布のなす空間 $(M, g, \nabla^{(\alpha)})$ は統計多様体であり，α–接続に関する曲率テンソルは

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

で表せる。ここで $c(\alpha) = (1-\alpha)(1+\alpha)$ である。

命題 5.5. Dirichlet 分布のなす空間の α–接続に関する曲率テンソルは (5.29) で与えられる。

次に式 (4.10), (4.11), (5.24), (5.25) から 1–接続及び (-1)–接続に関して平行となる概複素構造 $J^{(1)}$ 及び $J^{(-1)}$ の成分はそれぞれ $J_j^{(1)\,k} = P_j{}^k$ 及び

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

と表せることが分かる。これより

定理 5.6. Dirichlet 分布のなす空間において

(1) $(M, g, J^{(\pm 1)})$ はそれぞれ概 Hermite-like 多様体である。

(2) $(M, g, \nabla^{(\pm 1)}, J^{(\pm 1)})$ はそれぞれ Kähler-like 統計多様体である。

参　考　文　献

[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] K. Nomizu and T. Sasaki, *Affine Differential Geometry,* Cambridge Univ. Press, Cambridge, 1994.

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

[6] ———, Examples of the statistical submersion on the statistical model, Tensor N. S., 65 (2004), 170–178.

[7] ———, Statistical manifolds with almost contact structures and its statistical submersions, J. Geom., 85 (2006), 171–187.

[8] ———, Geodesics on statistical models of the multivariate normal distribution, Tensor N. S., 67 (2006), 162–169.

[9] ———, Examples of statistical manifolds with almost complex structures, preprint.

[10] ———, Exponential families admitting almost complex structures, preprint.

[11] M. Yamada, R. Ivanova, Y. Sasahara and T. Kawaguch, 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.

School of General Education
Shinshu University
Matsumoto 390–8621, Japan
E-mail address: ktakano@shinshu-u.ac.jp