振り返りと導入

前回は、指数型分布族にいくつかの構造を定め、Amari-Chentsov テンソルと α -接続を定義した。本稿では次のことを行う:

- 具体例の計算 (有限集合上の full support な確率分布の族)
- 具体例の計算 (正規分布族)

今回以降、次のように記法を変更する1)。

定義 0.1 (パラメータの空間の記法の変更). 可測空間 X 上の指数型分布族 $\mathcal P$ とその実現 (V,T,μ) に対し、

- 自然パラメータ空間 $\Theta_{(V,T,u)}$ を $\widetilde{\Theta}_{(V,T,u)}$ と書くことにし、
- 真パラメータ空間 $\Theta^{\mathcal{P}}_{(V,T,\mu)}$ を $\Theta_{(V,T,\mu)}$ と書くことにする。

文脈から明らかな場合は添字を省略することがある。

1 具体例: 有限集合上の full support な確率分布の族

本節では、有限集合上の full support な確率分布の族について、α-接続に関する測地線方程式を求めてみる。

設定 1.1 (有限集合上の full support な確率分布の族). $X := \{1, ..., n\} (n \in \mathbb{Z}_{\geq 1})$ とし、

$$\mathcal{P} := \left\{ \sum_{i=1}^{n} p_i \delta^i \in \mathcal{P}(X) \,\middle|\, 0 < p_i < 1 \,(i=1,\ldots,n) \right\} \tag{1.1}$$

とおく。ただし δ^i は 1 点 $i\in X$ での Dirac 測度である。 0425_資料 . pdf 例 3.1 でみたように $\mathcal P$ は X 上の指数型分布族である。

まず ρ が開であることを確かめる。

命題 1.2 (最小次元実現の構成およびP が開であることの確認).

(1) (V,T,μ) を次のように定めると、これは ρ の実現となる:

$$V = \mathbb{R}^{n-1},\tag{1.2}$$

$$T: X \to V, \quad k \mapsto {}^{t}(\delta_{1k}, \dots, \delta_{(n-1)k}),$$
 (1.3)

$$\mu = \gamma$$
 (数え上げ測度) (1.4)

- (2) この実現の対数分配関数 $\psi \colon \widetilde{\Theta} \to \mathbb{R}$ は $\psi(\theta) = \log \left(1 + \sum_{l=1}^{n-1} \exp \theta^l\right)$ となる。
- (3) 写像 $P := P_{(V,T,\mu)} : \widetilde{\Theta} \to \mathcal{P}$ は次をみたす:

$$P(\theta) = \frac{1}{1 + \sum_{l=1}^{n-1} \exp \theta^l} \left(\sum_{i=1}^{n-1} (\exp \theta^i) \delta^i + \delta^n \right)$$
 (1.5)

^{1) [}BN78] での記法によった。

(4) 次の写像 θ : $\mathcal{P} \to \Theta$ は P の逆写像である:

$$\theta \colon \mathcal{P} \to \Theta, \quad \sum_{i=1}^{n} p_i \delta^i \mapsto \left(\log \frac{p_i}{p_n} \right)_{i=1}^{n-1}$$
 (1.6)

- (5) $\Theta = \widetilde{\Theta} = \mathbb{R}^{n-1}$ が成り立つ。
- (6) (V,T,μ) は最小次元実現である。とくに \mathcal{P} は開である。

証明 (1) 実現であることはよい。

(2) 対数分配関数の定義より

$$\psi(\theta) = \log \int_{\mathcal{X}} \exp \langle \theta, T(x) \rangle \ \mu(dx) \tag{1.7}$$

$$= \log \sum_{i=1}^{n} \exp \left(\sum_{l=1}^{n-1} \theta^{l} \delta_{li} \right)$$
 (1.8)

$$= \log \left(\sum_{i=1}^{n-1} \exp \theta^i + 1 \right) \tag{1.9}$$

である。

- (3) $P(\theta) = \exp(\langle \theta, T(k) \rangle \psi(\theta)) \cdot \gamma$ を直接計算すれば得られる。
- (4) $P \circ \theta$, $\theta \circ P$ を直接計算すれば確かめられる。
- (5) すべての $\theta \in \mathbb{R}^{n-1}$ に対し $\int_{\mathcal{X}} \exp \langle \theta, T(x) \rangle \ \mu(dx) = \sum_{l=1}^{n-1} \exp \theta^l + 1 < \infty$ が成り立つから $\Theta = \mathbb{R}^{n-1}$ である。
- (6) 最小次元実現の特徴づけを確かめる。条件 A(3) と条件 B が成り立つことから、最小次元実現であることがわかる。

以降、 $\mathcal P$ には自然な位相および多様体構造が入っているものとして扱い、 $\mathcal P$ 上の自然な平坦アファイン接続を ∇ 、Fisher 計量を g とおく。また、 θ : $\mathcal P\to\Theta$ は多様体 $\mathcal P$ 上の座標とみなす。

図形的には、P は \mathbb{R}^{n-1} から \mathbb{R}^n 内の (n-1)-単体 (の内部) への微分同相写像になっている。

命題 1.3 (\mathcal{P} の 2 通りの位相と多様体構造). \mathcal{P} 上の位相 (resp. 多様体構造) として、 \mathcal{X} 上の符号付き測度全体のなすベクトル空間 $\mathcal{S}(\mathcal{X}) \cong \mathbb{R}^n$ の位相部分空間 (resp. 部分多様体) としての構造と、指数型分布族としての自然な構造の 2 通りを考えられるが、これらの構造は互いに一致する。

証明 いずれの位相 (resp. 多様体構造) に関しても写像 $P \colon \widetilde{\Theta} \to \mathcal{P}$ は同相 (resp. 微分同相) 写像だからである。

命題 1.4 (Fisher 計量の成分). 座標 $\theta = (\theta^1, \dots, \theta^{n-1})$ に関する Fisher 計量 g の成分は

$$g_{ij}(p) = \delta_{ij}p_i - p_ip_j \qquad (p \in \mathcal{P}, i, j = 1, ..., n - 1)$$
 (1.10)

となる。

証明 命題 2.2 の (V,T,μ) に関する Fisher 計量を \widetilde{g} とおくと、各点 $p \in \mathcal{P}$ に対し

$$g_{ij}(p) = g_p \left(\frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \theta^j} \right) \tag{1.11}$$

$$= (\theta^* \widetilde{g})_p \left(\frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \theta^j} \right) \tag{1.12}$$

$$=\widetilde{g}_{\theta(p)}\left(d\theta\left(\frac{\partial}{\partial\theta^{i}}\right),d\theta\left(\frac{\partial}{\partial\theta^{j}}\right)\right) \tag{1.13}$$

$$= (\operatorname{Hess} \psi)_{\theta(p)} \left(d\theta \left(\frac{\partial}{\partial \theta^{i}} \right), d\theta \left(\frac{\partial}{\partial \theta^{j}} \right) \right) \tag{1.14}$$

$$= (\operatorname{Var}_p[T])(e^i, e^j) \tag{1.15}$$

$$= E_p[(T^i - E_p[T^i])(T^j - E_p[T^j])]$$
(1.16)

$$= \sum_{l=1}^{n} (\delta_{il} - p_i)(\delta_{jl} - p_j)p_l$$
 (1.17)

$$=\delta_{ij}p_i - p_ip_j \tag{1.18}$$

が成り立つ。

命題 1.5 (AC テンソルの成分). 座標 θ に関する AC テンソル S の成分は

$$S_{ijk} = p_i \delta_{ij} \delta_{jk} - p_i p_k \delta_{ij} - p_i p_j \delta_{jk} - p_j p_k \delta_{ik} + 2p_i p_j p_k \qquad (p \in \mathcal{P}, i, j, k = 1, \dots, n - 1)$$

$$(1.19)$$

となる。

証明 前回 (0613_資料.pdf) の命題 1.9 を用いると

$$S_{ijk}(p) = E_v[(T^i - E_v[T^i])(T^j - E_v[T^j])(T^k - E_v[T^k])]$$
(1.20)

となるから、命題 1.4 と同様に直接計算して命題の主張の等式が得られる。

以降、n=3 の場合を考える。

命題 1.6 (n=3 での g,S,A の計算). 座標 θ に関し、g の行列表示は

$$(g_{ij})_{i,j} = \begin{pmatrix} p_1(1-p_1) & -p_1p_2 \\ -p_1p_2 & p_2(1-p_2) \end{pmatrix}, \quad (g^{ij})_{i,j} = \frac{1}{p_3} \begin{pmatrix} \frac{p_3}{p_1} + 1 & 1 \\ 1 & \frac{p_3}{p_2} + 1 \end{pmatrix}$$
(1.21)

となる。Sの成分は

$$S_{111} = p_1 - 3p_1^2 + 2p_1^3, (1.22)$$

$$S_{112} = S_{121} = S_{211} = -p_1 p_2 + 2p_1^2 p_2, (1.23)$$

$$S_{122} = S_{212} = S_{221} = -p_1 p_2 + 2p_1 p_2^2, (1.24)$$

$$S_{222} = p_2 - 3p_2^2 + 2p_2^3 (1.25)$$

となる。Aの成分は

$$A_{11}^{1} = 1 - 2p_1, A_{11}^{2} = 0 (1.26)$$

$$A_{11}^{1} = 1 - 2p_1,$$
 $A_{11}^{2} = 0$ (1.26)
 $A_{12}^{1} = A_{21}^{1} = -p_2,$ $A_{12}^{2} = A_{21}^{2} = -p_1$ (1.27)

$$A_{22}^{1} = 0,$$
 $A_{22}^{2} = 1 - 2p_2$ (1.28)

となる。

証明 g の行列表示は命題 1.4 よりわかる。その逆行列は直接計算よりわかる。S の成分は命題 1.5 よりわか る。Aの成分は直接計算よりわかる。

命題 1.7 (n=3 での測地線方程式). 各 $\alpha \in \mathbb{R}$ に対し、座標 θ に関する $\nabla^{(\alpha)}$ -測地線の方程式は

$$\ddot{\theta}^{1} = -\frac{1-\alpha}{2} \left((1-2p_1)\dot{\theta}^{1^2} - 2p_2\dot{\theta}^{1}\dot{\theta}^{2} \right)$$
 (1.29)

$$\ddot{\theta}^2 = -\frac{1-\alpha}{2} \left(-2p_1 \dot{\theta}^1 \dot{\theta}^2 + (1-2p_2) \dot{\theta}^2 \right)$$
 (1.30)

となる $(p_1, p_2$ は θ^1, θ^2 の関数であることに注意)。 とくに $\alpha=1$ のとき

$$\ddot{\theta}^1 = 0, \quad \ddot{\theta}^2 = 0 \tag{1.31}$$

である。

証明 測地線の方程式

$$\ddot{\theta}^{\dot{k}} = -\Gamma^{\dot{k}}_{i\dot{j}}\dot{\theta}^{\dot{i}}\dot{\theta}^{\dot{j}} \tag{1.32}$$

に、前回 (0613_資料.pdf) の命題 1.11 の等式 $\Gamma^{(\alpha)}_{ij} = \frac{1-\alpha}{2} A_{ij}^{k}$ を代入して得られる。

 $\alpha \neq 1$ の場合に上の測地線方程式を解くのは難しい (ように思う)。

具体例: 正規分布族

本節では、正規分布族について、α-接続に関する測地線方程式を求めてみる。

設定 2.1 (正規分布族). X := ℝ とし、

$$\mathcal{P} := \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \lambda(dx) \in \mathcal{P}(X) \,\middle|\, (\mu,\sigma) \in \mathbb{R} \times \mathbb{R}_{>0} \right\}$$
 (2.1)

とおく。

命題 2.2 (最小次元実現の構成およびP が開であることの確認).

(1) (V,T,μ) を次のように定めると、これは ρ の実現となる:

$$V = \mathbb{R}^2, \tag{2.2}$$

$$T: \mathcal{X} \to V, \quad x \mapsto {}^t(x, x^2),$$
 (2.3)

$$μ = λ$$
 (Lebesgue 測度). (2.4)

- (2) この実現の対数分配関数 $\psi \colon \widetilde{\Theta} \to \mathbb{R}$ は $\psi(\theta) = -\frac{(\theta^1)^2}{8\theta^2} \frac{1}{2}\log(-\theta^2) + \frac{1}{2}\log\pi$ となる。
- (3) 写像 $P := P_{(V,T,\mu)} : \widetilde{\Theta} \to \mathcal{P}$ は次をみたす:

$$P(\theta) = \exp\left(\theta^{1}x + \theta^{2}x^{2} - \psi(\theta)\right)\lambda(dx)$$
 (2.5)

(4) 次の写像 θ : $\mathcal{P} \to \Theta$ は P の逆写像である:

$$\theta: \mathcal{P} \to \Theta, \quad p \mapsto \left(\frac{E_p[x]}{\operatorname{Var}_p[x]}, -\frac{1}{2\operatorname{Var}_p[x]}\right)$$
 (2.6)

- (5) $\Theta = \widetilde{\Theta} = \mathbb{R} \times \mathbb{R}_{<0}$ が成り立つ。
- (6) (V,T,μ) は最小次元実現である。とくに \mathcal{P} は開である。

証明 (1) 実現であることはよい。

- (2) 直接計算よりわかる。
- (3) $P_{(V,T,\mu)}$ の定義よりわかる。
- (4) $(\theta^1, \theta^2) \in \mathbb{R} \times \mathbb{R}_{<0}$ と $(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_{>0}$ の対応に注意すれば直接計算よりわかる。
- (5) $\theta^2 \ge 0$ だと $\exp\left(\theta^1 x + \theta^2 x^2 \psi(\theta)\right)$ は積分可能でないから $\Theta \subset \widetilde{\Theta} \subset \mathbb{R} \times \mathbb{R}_{<0}$ である。逆に θ の定義より明らかに $\mathbb{R} \times \mathbb{R}_{<0} \subset \Theta$ である。したがって $\Theta = \widetilde{\Theta} = \mathbb{R} \times \mathbb{R}_{<0}$ である。
- (6) 最小次元実現の特徴づけの条件 A(3) と条件 B が成り立つことから、最小次元実現であることがわかる。

命題 2.3. Fisher 計量 g について

(1) 座標 (θ^1, θ^2) に関する g の成分は

$$g_{ij} = \left(-\frac{1}{2\theta^2}d\theta^1 + \frac{\theta^1}{2(\theta^2)^2}d\theta^2\right)d\theta^1 + \left(\frac{\theta^1}{2(\theta^2)^2}d\theta^1 + \left(\frac{1}{2(\theta^2)^2} - \frac{(\theta^1)^2}{2(\theta^2)^3}\right)d\theta^2\right)d\theta^2$$
(2.7)

である。

(2) 座標 (μ, σ) に関する g の成分は

$$g_{ij} = \frac{1}{\sigma^2} (d\mu)^2 + \frac{2}{\sigma^2} (d\sigma)^2$$
 (2.8)

である。

証明 座標 (θ^1, θ^2) と座標 (μ, σ) の間の座標変換は $\theta^1 = \frac{\mu}{\sigma^2}$, $\theta^2 = -\frac{1}{2\sigma^2}$ および $\mu = -\frac{\theta^1}{2\theta^2}$, $\sigma = \sqrt{-\frac{1}{2\theta^2}}$ である。

$$d\mu = -\frac{1}{2\theta^2}d\theta^1 + \frac{\theta^1}{2(\theta^2)^2}d\theta^2,$$
 (2.9)

$$d\sigma = \frac{1}{2\sqrt{2}}(-\theta^2)^{-3/2}d\theta^2,$$
(2.10)

$$d\theta^1 = \frac{1}{\sigma^2} d\mu - \frac{2\mu}{\sigma^3} d\sigma \tag{2.11}$$

$$d\theta^2 = \frac{1}{\sigma^3} d\sigma \tag{2.12}$$

よって

$$(d\theta^{1})^{2} = \frac{1}{\sigma^{4}}(d\mu)^{2} - \frac{\mu}{\sigma^{5}}d\mu d\sigma + \frac{4\mu^{2}}{\sigma^{6}}(d\sigma)^{2}$$
(2.13)

$$d\theta^1 d\theta^2 = \frac{1}{\sigma^5} d\mu d\sigma - \frac{2\mu}{\sigma^6} (d\sigma)^2 \tag{2.14}$$

$$(d\theta^2)^2 = \frac{1}{\sigma^6} (d\sigma)^2 \tag{2.15}$$

である。したがって

$$Dd\mu = \frac{1}{(\theta^2)^2} d\theta^1 d\theta^2 - \frac{\theta^1}{(\theta^2)^3} (d\theta^2)^2$$
 (2.16)

$$=\frac{4}{\sigma}d\mu d\sigma,\tag{2.17}$$

$$Dd\sigma = \frac{3}{4\sqrt{2}}(-\theta^2)^{-5/2}(d\theta^2)^2 \tag{2.18}$$

$$=\frac{3}{\sigma}(d\sigma)^2\tag{2.19}$$

である。よって

$$d\psi = \frac{\mu}{\sigma^2} d\mu + \left(-\frac{\mu^2}{\sigma^3} + \frac{1}{\sigma}\right) d\sigma \tag{2.20}$$

$$Hess \psi = Dd\psi \tag{2.21}$$

$$=d\left(\frac{\mu}{\sigma^2}\right)d\mu+\frac{\mu}{\sigma^2}Dd\mu+d\left(-\frac{\mu^2}{\sigma^3}+\frac{1}{\sigma}\right)d\sigma+\left(-\frac{\mu^2}{\sigma^3}+\frac{1}{\sigma}\right)Dd\sigma \tag{2.22}$$

$$= \frac{1}{\sigma^2} (d\mu)^2 + \frac{2}{\sigma^2} (d\sigma)^2 \tag{2.23}$$

である。座標変換により (θ^1, θ^2) に関する $\text{Hess} \psi$ の成分も得られる。

命題 2.4 (AC テンソルの成分). 座標 (μ, σ) に関する AC テンソル S の成分は

$$S_{111} = 0 (2.24)$$

$$S_{112} = S_{121} = S_{211} = \frac{2}{\sigma^3} \tag{2.25}$$

$$S_{122} = S_{212} = S_{221} = 0 (2.26)$$

$$S_{222} = \frac{8}{\sigma^3} \tag{2.27}$$

である。座標 (μ, σ) に関する A の成分は

$$A_{11}^{1} = 0, A_{11}^{2} = \frac{1}{\sigma}, (2.28)$$

$$A_{12}^{1} = A_{21}^{1} = \frac{2}{\sigma}, \qquad A_{12}^{2} = A_{21}^{2} = 0,$$
 (2.29)

$$A_{22}^{-1} = 0, A_{22}^{-2} = \frac{4}{\sigma} (2.30)$$

である。

証明

$$DD\psi = D\left(\frac{1}{\sigma^2}(d\mu)^2 + \frac{2}{\sigma^2}(d\sigma)^2\right)$$
 (2.31)

$$= -\frac{2}{\sigma^3} (d\mu)^2 d\sigma + \frac{1}{\sigma^2} D(d\mu)^2 - \frac{4}{\sigma^3} (d\sigma)^3 + \frac{2}{\sigma^2} D(d\sigma)^2$$
 (2.32)

ここで

$$D(d\mu)^2 = 2d\mu Dd\mu = \frac{8}{\sigma}(d\mu)^2 d\sigma \tag{2.33}$$

$$D(d\sigma)^2 = 2d\sigma Dd\sigma = \frac{6}{\sigma}(d\sigma)^3$$
 (2.34)

だから

$$DDd\psi = \frac{6}{\sigma^3} (d\mu)^2 d\sigma + \frac{8}{\sigma^3} (d\sigma)^3$$
 (2.35)

である。これより命題の主張の式が得られる。Aの成分は直接計算より得られる。

命題 2.5 (接続係数).

(1) 座標 (μ, σ) に関する Γ^g の成分は

$$\Gamma_{11}^{g_{11}} = 0, \qquad \Gamma_{12}^{g_{12}} = \Gamma_{21}^{g_{21}} = -\frac{1}{\sigma}, \qquad \Gamma_{22}^{g_{22}} = 0,$$
(2.36)

$$\Gamma^{g_{11}^2} = \frac{1}{2\sigma}, \qquad \Gamma^{g_{12}^2} = \Gamma^{g_{21}^2} = 0, \qquad \Gamma^{g_{22}^2} = -\frac{1}{\sigma}$$
(2.37)

である。

(2) 座標 (μ, σ) に関する $\Gamma^{(\alpha)}$ $(\alpha \in \mathbb{R})$ の成分は

$$\Gamma^{(\alpha)}_{11}^{1} = 0, \qquad \Gamma^{(\alpha)}_{12}^{1} = \Gamma^{(\alpha)}_{21}^{1} = -\frac{1+\alpha}{\sigma}, \qquad \Gamma^{(\alpha)}_{22}^{1} = 0,$$
 (2.38)

$$\Gamma^{(\alpha)}_{11}^2 = \frac{1-\alpha}{2\sigma}, \qquad \Gamma^{(\alpha)}_{12}^2 = \Gamma^{(\alpha)}_{21}^2 = 0, \qquad \qquad \Gamma^{(\alpha)}_{22}^2 = -\frac{1+2\alpha}{\sigma}$$
 (2.39)

である。

証明 Γ^g は $\Gamma^g_{ij}^k = \frac{1}{2} g^{kl} \left(\partial_i g_{jl} + \partial_j g_{li} - \partial_l g_{ij} \right)$ を直接計算することで得られる。 $\Gamma^{(\alpha)}$ は $\Gamma^{(\alpha)}_{ij}^k = \Gamma^g_{ij}^k - \frac{\alpha}{2} A_{ij}^k$ より得られる。

命題 2.6 (測地線方程式). (μ, σ) 座標に関する $\nabla^{(\alpha)}$ -測地線の方程式は

$$\begin{cases}
\ddot{\mu} - \frac{2(1+\alpha)}{\sigma}\dot{\mu}\dot{\sigma} = 0, \\
\ddot{\sigma} + \frac{1-\alpha}{2\sigma}\dot{\mu}^2 - \frac{1+2\alpha}{\sigma}\dot{\sigma}^2 = 0
\end{cases}$$
(2.40)

である。とくに $\alpha = 0$ のとき

$$\begin{cases}
\ddot{\mu} - \frac{2}{\sigma}\dot{\mu}\dot{\sigma} = 0, \\
\ddot{\sigma} + \frac{1}{2\sigma}\dot{\mu}^2 - \frac{1}{\sigma}\dot{\sigma}^2 = 0
\end{cases}$$
(2.41)

である。

証明 測地線の方程式「 $\ddot{x^k} = -\Gamma^k_{ij}\dot{x^i}\dot{x^j}$ 」に接続係数を代入して得られる。

事実 2.7. $\mathbb{H} := \mathbb{R} \times \mathbb{R}_{>0}$ とし、 \S を \mathbb{H} 上の双曲計量とする。このとき、写像 $f : \mathbb{H} \to \mathcal{P}$, $(x,y) \to (x,\sqrt{2}y)$ は (\mathbb{H},\S) から (\mathcal{P},g) への等長同型写像である。

事実 2.8. 局所等長同型写像は測地線を保つ。

事実 2.9. $\mathbb H$ の測地線は、x 軸上に中心を持つ円弧または y 軸に平行な直線である。

命題 2.10. 𝔻 の測地線は、長軸が x 軸に重なる 長軸: 短軸 = 2:1 の楕円または y 軸に平行な直線である。

今後の予定

[TODO]

参考文献

[Ama16] Shun-ichi Amari, **Information Geometry and Its Applications**, Applied Mathematical Sciences, vol. 194, Springer Japan, Tokyo, 2016 (en).

[BN78] O. E. Barndorff-Nielsen, Information and exponential families: In statistical theory, Wiley, 1978.

A 付録

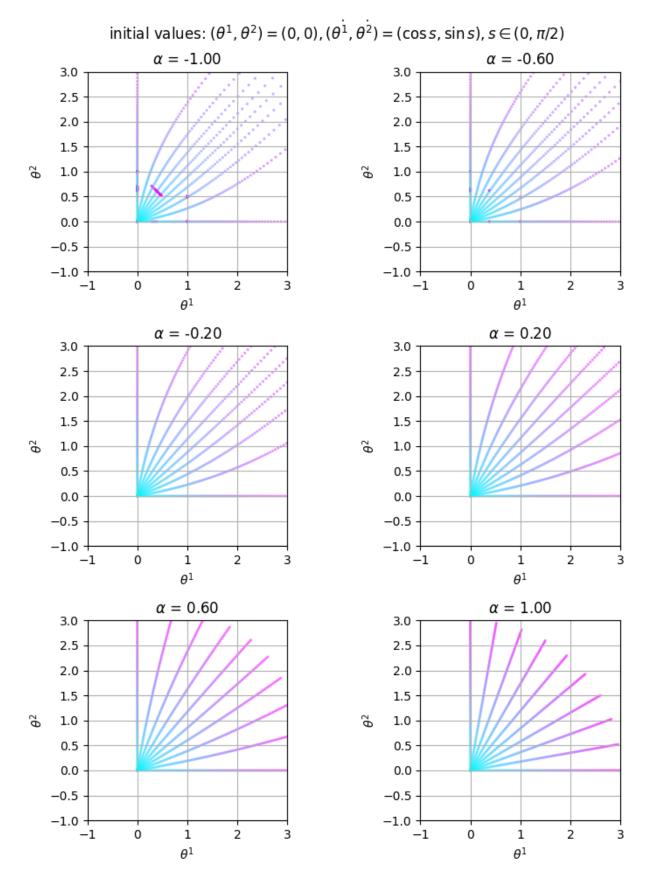


図 1 α を変化させたときの ∇^{α} -測地線の様子