1 振り返りと導入

- 指数型分布族の "多様体化"
- Amari-Chentsov テンソル

2 指数型分布族の "多様体化"

定義 2.1 (自然パラメータ付け).

$$P\colon \Theta_{(V,T,\mu)} \to \mathcal{P}(\mathcal{X}), \quad \theta \mapsto \exp(\langle \theta, T(x)\rangle - \psi(\theta))\,\mu(dx) \tag{2.1}$$

を (V,T,μ) による $\mathcal{P}(\mathcal{X})$ の**自然パラメータ付け (natural parameterization)** という。

定義 2.2 (真パラメータ空間).

$$\Theta^{\mathcal{P}}_{(V,T,\mu)} := P^{-1}(\mathcal{P}) \tag{2.2}$$

を $\mathcal{P}$ の (V,T,μ) に関する**真パラメータ空間 (strict parameter space)** という。

A. 最小次元実現

命題-定義 2.3 (単射性条件). $\mathcal{P}$ の実現 (V,T,μ) に関する次の条件は同値である:

(1) $P\colon \Theta_{(V,T,\mu)} \to \mathcal{P}(\mathcal{X})$ は単射である。
(2) $\forall\theta \in V^\vee$ に対し「$\langle \theta, T(x)\rangle = \text{const.}\ \mu\text{-a.e.}x \implies \theta = 0$」が成り立つ。
(3) V の任意の真アファイン部分空間 W に対し、「$T(x) \in W\ \mu\text{-a.e.}x$ でない」が成り立つ。

これらの条件が成り立つとき、(V,T,μ) は**単射性条件**をみたすという。

証明 [TODO] □

命題-定義 2.4 (Affine span 条件). $\mathcal{P}$ の実現 (V,T,μ) に関する条件

(1) $\Theta^{\mathcal{P}}_{(V,T,\mu)}$ は $V^\vee$ を affine span する。

が成り立つとき、(V,T,μ) は **affine span 条件**をみたすという。[TODO] 違う名称にすべき?

注意 2.5 (条件をみたさない例). [TODO]

命題 2.6 (実現の間の変換). $(V,T,\mu),(V',T',\mu')$ を $\mathcal{P}$ の実現、P,P' をそれぞれの実現による自然パラメータ付けとする。このとき、(V,T,μ) が単射性条件と affine span 条件をみたすならば、ある全射 $\mathbb{R}$-線型写像 $L\colon V' \to V$ とベクトル $b \in V$ が存在して、すべての $p \in \mathcal{P}$ に対し

$$T(x) = L(T'(x)) + b \qquad p\text{-a.e.}x \tag{2.3}$$

が成り立つ。[TODO] μ-a.e. にすべき?

証明 P, P' の右逆写像 $\theta\colon \mathcal{P} \to \Theta^{\mathcal{P}}_{(V,T,\mu)}$ および $\theta'\colon \mathcal{P} \to \Theta^{\mathcal{P}}_{(V',T',\mu')}$ をひとつずつ選んでおく。

まず $\mu = \mu'$ の場合を考える。

まず (V, T, μ) の affine span 条件より、ある $a^i \in \Theta^{\mathcal{P}}_{(V,T,\mu)}$ $(i = 0, \ldots, m)$ が存在して、$e^i := a^i - a^0$ $(i = 1, \ldots, m)$ は $V^\vee$ の基底となる。そこで $p^i := P(a^i) \in \mathcal{P}$ $(i = 0, \ldots, m)$ とおき、線型写像 $L\colon V' \to V$ を $t' \mapsto \left\langle \theta'(p^i) - \theta'(p^0), t' \right\rangle e_i$ で定め、ベクトル $b \in V$ を $b := \left\{ \psi(\theta(p^i)) - \psi(\theta(p^i)) - \psi(\theta'(p^0)) + \psi(\theta'(p^0)) \right\} e_i$ で定める。すべての $p \in \mathcal{P}$ に対し

$$T(x) = L(T'(x)) + b \quad p\text{-a.e.}x \tag{2.4}$$

が成り立つことを示す。

指数型分布族の定義より、すべての $p \in \mathcal{P}$ に対し

$$\exp(\langle \theta(p), T(x) \rangle - \psi(\theta(p))) = \frac{dp}{d\mu}(x) = \exp(\langle \theta'(p), T'(x) \rangle - \psi'(\theta'(p))) \qquad p\text{-a.e.}x \tag{2.5}$$

が成り立つ $(p, \mu, \mu'$ が互いに絶対連続であることを用いた)。このことを「a.e.」を使わずに言い換えると次のようになる: すべての $p \in \mathcal{P}$ に対し、ある p-零集合 $N \subset \mathcal{X}$ が存在して、すべての $x \in \mathcal{X} \setminus N$ に対し上の等式が成り立つ。$p \in \mathcal{P}$ とする。上の等式を整理して

$$\langle \theta(p), T(x) \rangle - \langle \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) \qquad (x \in \mathcal{X} \setminus N) \tag{2.6}$$

が成り立つ。そこで $x_0 \in \mathcal{X} \setminus N$ をひとつ選んで固定すると

$$\langle \theta(p), T(x) - T(x_0) \rangle = \langle \theta'(p), T'(x) - T'(x_0) \rangle \qquad (x \in \mathcal{X} \setminus N) \tag{2.7}$$

が成り立つ。

各 $i = 0, \ldots, m$ に対し

$$\left\langle a^i, T(x) - T(x_0) \right\rangle = \left\langle \theta'(p^i), T'(x) - T'(x_0) \right\rangle \qquad (x \in \mathcal{X} \setminus N) \tag{2.8}$$

となるから、各 $i = 1, \ldots, m$ に対し

$$\begin{aligned}
\left\langle e^i, T(x) - T(x_0) \right\rangle &= \left\langle \theta'(p^i) - \theta'(p^0), T'(x) - T'(x_0) \right\rangle \qquad (x \in \mathcal{X} \setminus N) && (2.9)\\
\therefore \quad \left\langle e^i, T(x) \right\rangle &= \left\langle \theta'(p^i) - \theta'(p^0), T'(x) \right\rangle && (2.10)\\
&\quad - \left\langle \theta'(p^i) - \theta'(p^0), T'(x_0) \right\rangle + \left\langle e^i, T(x_0) \right\rangle \qquad (x \in \mathcal{X} \setminus N) && (2.11)
\end{aligned}$$

が成り立つ。よって、すべての $x \in \mathcal{X} \setminus N$ に対し

$$\begin{aligned}
T(x) &= L(T'(x)) + \left\{ \left\langle \theta'(p^i) - \theta'(p^0), T'(x_0) \right\rangle + \left\langle e^i, T(x_0) \right\rangle \right\} e_i && (2.12)\\
&= L(T'(x)) + b && (2.13)
\end{aligned}$$

が成り立つ。もし L が全射でなかったとすると、$T(x) = L(T'(x)) + b \in \operatorname{Im} L + b$ が p-a.e.x すなわち μ-a.e.x に対し成り立つことになるが、$\operatorname{Im} L + b$ は V の真アファイン部分空間だから (V, T, μ) の単射性条件に反する。したがって L は全射である。

$\mu \neq \mu'$ の場合も、(V, T, μ) の単射性条件を用いて $(V, T, \mu') \overset{\mathrm{id},0}{\rightsquigarrow} (V, T, \mu)$ が成り立つことから明らか。 □

系 2.7. 上の命題の状況でさらに (V', T', μ') も単射性条件と affine span 条件をみたすならば、L は線型同型であり、すべての $p \in \mathcal{P}$ に対し

$$\begin{cases} T(x) = L(T'(x)) + b & p\text{-a.e.}x \\ \theta'(p) = {}^t L(\theta(p)) \end{cases} \tag{2.14}$$

が成り立つ。ただし写像 $\theta \colon \mathcal{P} \to \Theta^{\mathcal{P}}_{(V,T,\mu)}$ および $\theta' \colon \mathcal{P} \to \Theta^{\mathcal{P}}_{(V',T',\mu')}$ は P, P' の逆写像である。[TODO] L は一意?

証明 上の証明の続きで、すべての $x \in \mathcal{X} \setminus N$ に対し

$$\langle \theta(p), L(T'(x)) + b \rangle - \langle \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) \tag{2.15}$$

$$\therefore \quad \langle {}^t L(\theta(p)), T'(x) \rangle - \langle \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) - \langle \theta(p), b \rangle \tag{2.16}$$

$$\therefore \quad \langle {}^t L(\theta(p)) - \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) - \langle \theta(p), b \rangle \tag{2.17}$$

が成り立つ。(V', T', μ) の単射性条件より ${}^t L(\theta(p)) = \theta'(p)$ が成り立つ。

θ, θ' の役割を入れ替えて ${}^t L(\theta'(p)) = \theta(p)$ も成り立つから、とくに $V^\vee$ の基底 e^i $(i = 1, \ldots, m)$ に対し ${}^t L\, {}^t L e^i = e^i$ が成り立つ。よって ${}^t L, L$ は線型同型写像である。 □

定理 2.8 (最小次元実現の特徴づけ)**.** $\mathcal{P}$ の実現 (V, T, μ) に関する次の条件は同値である:

(1) (V, T, μ) は $\mathcal{P}$ の最小次元実現である。
(2) (V, T, μ) は単射性条件と affine span 条件をみたす。

証明 (1) ⇒(2) 以前示した。[TODO] affine span 条件は示してない?

(2) ⇒(1) (V, T, μ) が単射性条件と affine span 条件をみたすとする。$\mathcal{P}$ の任意の実現 (V', T', μ') に対し、命題 2.6 より全射線型写像 $L : V' \to V$ が存在するから、$\dim V \le \dim V'$ である。したがって V は $\mathcal{P}$ の最小次元実現である。 □

注意 2.9 (正規分布族の最小次元実現)**.** [TODO]

命題-定義 2.10. 指数型分布族 $\mathcal{P}$ に関し、次は同値である:

(1) ある最小次元実現 (V, T, μ) に対し、$\Theta^{\mathcal{P}}_{(V,T,\mu)}$ は $V^\vee$ で開である。
(2) すべての最小次元実現 (V, T, μ) に対し、$\Theta^{\mathcal{P}}_{(V,T,\mu)}$ は $V^\vee$ で開である。

$\mathcal{P}$ がこれらの同値な 2 条件をみたすとき、$\mathcal{P}$ は**開 (open)** であるという。

証明 (1) ⇒ (2) は命題 2.6 より従う。(2) ⇒ (1) は最小次元実現が存在することから従う。 □

B. 多様体構造

命題-定義 2.11. $\mathcal{P}$ は開であるとする。$\mathcal{P}$ の最小次元実現 (V,T,μ) をひとつ選ぶと、自然パラメータ付け P により、$\mathcal{P}$ 上に多様体構造と平坦アファイン接続を定めることができる。この多様体構造および平坦アファイン接続は最小次元実現のとり方によらない。これを $\mathcal{P}$ の**自然な多様体構造**および**自然な平坦アファイン接続**と呼ぶ。

証明 下の図式の可換性と、P^{-1}, P'^{-1} が diffeo. であること、${}^tL$ が線型同型であることから従う。

$$\begin{tikzcd} (\mathcal{P}, \mathcal{U}_{(V,T,\mu)}) \arrow[r, "\mathrm{id}"] \arrow[d, "P^{-1}"'] & (\mathcal{P}, \mathcal{U}_{(V',T',\mu')}) \arrow[d, "P'^{-1}"] \\ \Theta^{\mathcal{P}} \arrow[r, "{}^tL"] & \Theta'^{\mathcal{P}} \end{tikzcd} \tag{2.18}$$

□

命題-定義 2.12 ($\mathcal{P}$ の Fisher 計量). $\mathcal{P}$ は開であるとする。$\mathcal{P}$ の最小次元実現 (V,T,μ) をひとつ選ぶと、$\Theta^{\mathcal{P}}_{(V,T,\mu)}$ 上の Fisher 計量 g を θ で引き戻して $\mathcal{P}$ 上の Riemann 計量 $\theta^* g$ が定まる。この計量は最小次元実現のとり方によらない。これを $\mathcal{P}$ 上の **Fisher 計量**と呼ぶ。

証明 期待値と分散のペアリングの命題と同様の議論により、各 $p \in \mathcal{P}$ に対し $g_{\theta(p)} = (L \otimes L) g'_{\theta'(p)}$ が成り立つ。

示すべきことは $\theta^* g = \theta'^* g'$ が成り立つことである。各 $p \in \mathcal{P}$, $u, v \in T_p\mathcal{P}$ に対し

$$\begin{aligned}
(\theta^* g)_p(u,v) &= g_{\theta(p)}(d\theta_p(u), d\theta_p(v)) && (2.19)\\
&= \left\langle g_{\theta(p)}, d\theta_p(u) \otimes d\theta_p(v) \right\rangle && (2.20)\\
&= \left\langle (L \otimes L) g'_{\theta'(p)}, d\theta_p(u) \otimes d\theta_p(v) \right\rangle && (2.21)\\
&= \left\langle g'_{\theta'(p)}, {}^tL \circ d\theta_p(u) \otimes {}^tL \circ d\theta_p(v) \right\rangle && (2.22)\\
&= \left\langle g'_{\theta'(p)}, d({}^tL \circ \theta)_p(u) \otimes d({}^tL \circ \theta)_p(v) \right\rangle && (2.23)\\
&= \left\langle g'_{\theta'(p)}, d\theta'_p(u) \otimes d\theta'_p(v) \right\rangle && (2.24)\\
&= g'_p(d\theta'_p(u), d\theta'_p(v)) && (2.25)\\
&= (\theta'^* g')_p(u,v) && (2.26)
\end{aligned}$$

が成り立つ。 □

[TODO] Amari-Chentsov もいける?

3 期待値パラメータ空間

定義 3.1 (期待値パラメータ空間). 集合 $\mathcal{M}_{(V,T,\nu)}$

$$\mathcal{M}_{(V,T,\nu)} := \left\{\mu \in V \mid \exists\, p : \mathcal{X} \text{ 上の確率分布 s.t. } p \ll \nu,\ E_p[T] = \mu\right\} \tag{3.1}$$

を (V, T, ν) の**期待値パラメータ空間 (mean parameter space)** という。

期待値パラメータ空間 $\mathcal{M}$ は、$\mathcal{P}$ に属する確率分布に関する T の期待値をすべて含んでいる (一般には真に含んでいる)。

命題 3.2. $\mu \in V$ がある $p \in \mathcal{P}$ に関する T の期待値ならば (すなわち $\mu = E_p[T]$ ならば)、μ は $\mathcal{M}_{(V,T,\nu)}$ に属する。

証明 [TODO] □

命題 3.3 ($\mathcal{M}$ は凸集合). $\mathcal{M}_{(V,T,\nu)}$ は V の凸集合である。

証明 [TODO] □

4 Fisher 計量

例 4.1 (正規分布族). [TODO] ちゃんと書く $\mathcal{P}$ を $\mathcal{X} = \mathbb{R}$ 上の正規分布族とし、実現 (V, T, μ) を $V = \mathbb{R}^2$, $T(x) = (x, x^2)$, $\mu = \lambda$ とおく。これは条件 A をみたす。

自然パラメータ空間は $\Theta = \Theta^\circ = \mathbb{R} \times \mathbb{R}_{<0}$ である。

対数分配関数は

$$\psi(\theta) = \frac{\mu^2}{2\sigma^2} + \log\sigma + \frac{1}{2}\log 2\pi \tag{4.1}$$

である。ただし $\theta^1 = \frac{\mu}{\sigma^2}$, $\theta^2 = -\frac{1}{2\sigma^2}$ とおいた。よって

$$\begin{aligned} d\psi &= \frac{\mu}{\sigma^2}d\mu + \frac{\sigma^2-\mu^2}{\sigma^3}d\sigma && (4.2)\\ &= -\frac{\theta^1}{2\theta^2}d\theta^1 + \left(-\frac{1}{2\theta^2} + \frac{(\theta^1)^2}{4(\theta^2)^2}\right)d\theta^2 && (4.3)\\ \operatorname{Hess}\psi &= Dd\psi && (4.4)\\ &= \left(-\frac{1}{2\theta^2}d\theta^1 + \frac{\theta^1}{2(\theta^2)^2}d\theta^2\right)d\theta^1 + \left(\frac{\theta^1}{2(\theta^2)^2}d\theta^1 + \left(\frac{1}{2(\theta^2)^2} - \frac{(\theta^1)^2}{2(\theta^2)^3}\right)d\theta^2\right)d\theta^2 && (4.5)\\ &= \frac{1}{\sigma^2}(d\mu)^2 + \frac{2}{\sigma^2}(d\sigma)^2 && (4.6)\end{aligned}$$

である。Fisher 計量 $g := \operatorname{Hess}\psi$ から定まる Levi-Civita 接続 $\nabla^{(g)}$ の、座標 μ, σ に関する接続係数を求めてみる。

$$\begin{aligned} {\Gamma^{(g)}}^1_{11} &= 0, & {\Gamma^{(g)}}^1_{12} &= {\Gamma^{(g)}}^1_{21} = -\frac{1}{\sigma}, & {\Gamma^{(g)}}^1_{22} &= 0, && (4.7)\\ {\Gamma^{(g)}}^2_{11} &= \frac{1}{2\sigma}, & {\Gamma^{(g)}}^2_{12} &= {\Gamma^{(g)}}^2_{21} = 0, & {\Gamma^{(g)}}^2_{22} &= -\frac{1}{\sigma} && (4.8)\end{aligned}$$

測地線の方程式は

$$\begin{cases} x'' - \dfrac{2}{y}x'y' = 0 \\ y'' + \dfrac{1}{2y}(x')^2 - \dfrac{1}{y}(y')^2 = 0 \end{cases} \tag{4.9}$$

である。これを直接解くのは少し大変である。その代わりに、既知の Riemann 多様体との間の等長同型を利用して測地線を求める。(Θ, g) は、上半平面 H に計量 $\breve{g} = \dfrac{(dx)^2 + (dy)^2}{2y^2}$ を入れた Riemann 多様体との間に等長同型 $(\Theta, g) \to (H, \breve{g})$, $(x, y) \mapsto (x, \sqrt{2}y)$ を持つ。Levi-Civita 接続に関する測地線は等長同型で保たれるから、$(H, \breve{g})$ の測地線を求めればよい。$(H, \breve{g})$ の測地線は、y 軸に平行な直線と x 軸上に中心を持つ半円で尽くされることが知られている。これらを等長同型で写して、(Θ, g) の測地線として y 軸に平行な直線と x 軸上に長軸を持つ半楕円が得られる。

5 今後の予定

- KL ダイバージェンス
- Fisher 計量
- アファイン接続

6 参考文献

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