振り返りと導入

前回は自然パラメータ空間に Fisher 計量を定義した。本稿では次のことを行う:

- 最小次元実現の間のアファイン変換の一意存在を述べる。

$\mathcal{X}$ を可測空間、$\mathcal{P} \subset \mathcal{P}(\mathcal{X})$ を $\mathcal{X}$ 上の指数型分布族とする。新たな用語として次の 2 つを導入しておく。

定義 0.1 (自然パラメータ付け). (V, T, μ) を $\mathcal{P}$ の実現、ψ を (V, T, μ) の対数分配関数とする。

$$P_{(V,T,\mu)} \colon \Theta_{(V,T,\mu)} \to \mathcal{P}(\mathcal{X}), \quad \theta \mapsto \exp(\langle \theta, T(x) \rangle - \psi(\theta)) \cdot \mu \tag{0.1}$$

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

定義 0.2 (真パラメータ空間). (V, T, μ) を $\mathcal{P}$ の実現とする。

$$\Theta^{\mathcal{P}}_{(V,T,\mu)} \coloneqq P^{-1}_{(V,T,\mu)}(\mathcal{P}) \tag{0.2}$$

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。

以降、$\mathcal{P}$ の実現 $(V, T, \mu), (V', T', \mu')$ に対し、それぞれによる自然パラメータ付けを P, P' と略記したり、それぞれに関する真パラメータ空間を Θ, Θ' と略記したりすることがある。

1 最小次元実現の間のアファイン変換

本節の目標は、最小次元実現の間のアファイン変換の一意存在を述べた定理 1.13 の証明である。本節では、ステートメントを簡潔にするために圏の言葉を用いる。

命題-定義 1.1 ($\mathcal{P}$ の実現とアファイン写像の圏). 次のデータにより圏が定まる:

- 対象: $\mathcal{P}$ の実現 (V, T, μ) 全体
- 射: (V, T, μ) から (V', T', μ') への射は、V から V' への全射アファイン写像 (L, b) $(L \in \mathrm{Lin}(V, V'),\ b \in V')$ であって $T'(x) = L(T(x)) + b$ μ-a.e.x をみたすもの
- 合成: アファイン写像の合成 $(L, b) \circ (K, c) = (LK, Lc + b)$

この圏を **$\mathcal{P}$ の実現とアファイン写像の圏**と呼び、$\mathbf{C}_{\mathcal{P}}$ と書く。

証明 示すべきことは、射の合成が射であること、恒等射の存在、結合律の 3 点である。射の合成が射であることは、全射と全射の合成が全射であることと、$\mu \ll \mu'$ かつ $\mu' \ll \mu$ が成り立つことより従う。また、(V, T, μ) の恒等射は明らかに恒等写像 $(\mathrm{id}_V, 0)$ であり、結合律はアファイン写像の合成の結合律より従う。 □

最小次元実現を特徴づける 2 つの条件を導入する。

命題-定義 1.2 (条件 A). $\mathcal{P}$ の実現 (V, T, μ) に関する次の条件は同値である:

(1) $P \colon \Theta \to \mathcal{P}(\mathcal{X})$ は単射である。

(2) $\forall\theta \in V^\vee$ に対し「$\langle \theta, T(x)\rangle = \mathrm{const.}\ \mu\text{-a.e.}x \implies \theta = 0$」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**条件 A** をみたすという。

証明 (1) $\iff$ (2) は `0502_資料.pdf` の命題 2.2 で示した。(2) $\iff$ (3) は `0523_コメント.pdf` の命題 0.4 に記した。 □

定義 1.3 (条件 B). $\mathcal{P}$ の実現 (V, T, μ) に関する条件

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

が成り立つとき、(V, T, μ) は**条件 B** をみたすという。

条件 A は射の一意性を保証する。

補題 1.4 (射の一意性). $(V, T, \mu), (V', T', \mu')$ を $\mathbf{C}_{\mathcal{P}}$ の対象とする。このとき、(V, T, μ) が条件 A をみたすならば (V, T, μ) から (V', T', μ') への射は一意である。

証明 $(L, b), (K, c)$ を (V, T, μ) から (V', T', μ') への射とする。射の定義より

$$\begin{cases} T'(x) = L(T(x)) + b & \mu\text{-a.e.}x \\ T'(x) = K(T(x)) + c & \mu\text{-a.e.}x \end{cases} \tag{1.1}$$

が成り立つから、2 式を合わせて

$$(K - L)(T(x)) = b - c \qquad \mu\text{-a.e.}x \tag{1.2}$$

となる。そこで基底を固定して成分ごとに (V, T, μ) の条件 A(2) を適用すれば、$K = L$ を得る。よって上式で $K = L$ として $b = c$ μ-a.e. したがって $b = c$ を得る。以上より $(L, b) = (K, c)$ である。 □

射が存在するための十分条件を調べる。この後の中間目標は補題 1.7 であり、その証明は次の図のような 2 段階に分けて行う。

$$\begin{tikzcd}
(V', T', \mu') \arrow[rr, "\text{補題 1.7}"] \arrow[dr, "\text{補題 1.5}"'] & & (V, T, \mu) \\
& (V', T', \mu) \arrow[ur, "\text{補題 1.6}"'] &
\end{tikzcd} \tag{1.3}$$

補題 1.5 (基底測度だけを変化させた対象からの射). (V, T, μ') から (V, T, μ) への射が存在する。

証明 $(\mathrm{id}, 0)$ が求める射である。 □

補題 1.6 (基底測度以外を変化させた対象からの射). (V, T, μ) が 条件 A と条件 B をみたすならば、任意の対象 (V', T', μ) から (V, T, μ) への射が存在する。

証明 Step 0: $V, V^\vee$ の基底を選ぶ　(V, T, μ) の条件 B より、$V^\vee$ のあるアファイン基底 $a^i \in \Theta^{\mathcal{P}}$ $(i = 0, \ldots, m)$ が存在する。そこで $e^i := a^i - a^0 \in V^\vee$ $(i = 1, \ldots, m)$ とおくとこれは $V^\vee$ の基底である。さらに e^i の双対基底を V の元と同一視したものを $e_i \in V$ $(i = 1, \ldots, m)$ とおいておく。

Step 1: 射 (L, b) の構成　P, P' の右逆写像 $\theta\colon \mathcal{P} \to \Theta^{\mathcal{P}}$ および $\theta'\colon \mathcal{P} \to \Theta'^{\mathcal{P}}$ をひとつずつ選んで $p^i := P(a^i) \in \mathcal{P}$ $(i = 0, \ldots, m)$ とおき、(L, b) を次のように定める:

$$L\colon V' \to V, \quad t' \mapsto \left\langle \theta'(p^i) - \theta'(p^0), t' \right\rangle e_i \tag{1.4}$$

$$b := \left\{ \psi(\theta(p^i)) - \psi(\theta(p^i)) - \psi(\theta'(p^0)) + \psi(\theta'(p^0)) \right\} e_i \in V \tag{1.5}$$

示すべきことは、すべての $p \in \mathcal{P}$ に対し

$$T(x) = L(T'(x)) + b \quad \mu\text{-a.e.}x \tag{1.6}$$

が成り立つことと、(L, b) が全射となることである。

Step 2: $T(x) = L(T'(x)) + b$ の証明　指数型分布族の定義より、任意の $p \in \mathcal{P}$ に対し、ある μ-零集合 $N_p \subset \mathcal{X}$ が存在して

$$\exp(\langle \theta(p), T(x) \rangle - \psi(\theta(p))) = \frac{dp}{d\mu}(x) = \exp(\langle \theta'(p), T'(x) \rangle - \psi'(\theta'(p))) \qquad (x \in \mathcal{X} \setminus N_p) \tag{1.7}$$

$$\therefore \qquad \langle \theta(p), T(x) \rangle - \langle \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) \qquad (x \in \mathcal{X} \setminus N_p) \tag{1.8}$$

が成り立つ。とくに各 $i = 0, \ldots, m$ に対し

$$\left\langle \theta(p^i), T(x) \right\rangle - \left\langle \theta'(p^i), T'(x) \right\rangle = \psi(\theta(p^i)) - \psi'(\theta'(p^i)) \qquad (x \in \mathcal{X} \setminus N_{p^i}) \tag{1.9}$$

が成り立つから、(i の場合) − (0 の場合) を考えて

$$\begin{aligned} &\left\langle \theta(p^i) - \theta(p^0), T(x) \right\rangle - \left\langle \theta'(p^i) - \theta'(p^0), T'(x) \right\rangle \\ &= \psi(\theta(p^i)) - \psi'(\theta'(p^i)) - \psi(\theta(p^0)) + \psi'(\theta'(p^0)) \qquad (x \in \mathcal{X} \setminus (N_{p^0} \cup N_{p^i})) \end{aligned} \tag{1.10}$$

となる。ここで (V, T, μ) の条件 A (1) より $\theta(p^i) = a^i$ が成り立つから、(1.10) より

$$\begin{aligned} \left\langle a^i - a^0, T(x) \right\rangle = &\left\langle \theta'(p^i) - \theta'(p^0), T'(x) \right\rangle \\ &+ \psi(\theta(p^i)) - \psi'(\theta'(p^i)) - \psi(\theta(p^0)) + \psi'(\theta'(p^0)) \qquad (x \in \mathcal{X} \setminus (N_{p^0} \cup N_{p^i})) \end{aligned} \tag{1.11}$$

したがって

$$T(x) = L(T'(x)) + b \qquad (x \in \mathcal{X} \setminus (N_{p^0} \cup \cdots \cup N_{p^m})) \tag{1.12}$$

が成り立つ。これで Step 2 が完了した。

Step 3: (L, b) が全射であることの証明　L が全射であることをいえばよい。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) の条件 A (3) に反する。したがって L は全射である。 □

補題 1.7 (2 条件をみたす対象への射). (V, T, μ) が条件 A と条件 B をみたすならば、任意の対象 (V', T', μ') から (V, T, μ) への射が存在する。

証明 上の 2 つの補題より存在する 2 つの射 $(V', T', \mu') \to (V', T', \mu) \to (V, T, \mu)$ を合成すればよい。 □

各条件をみたさない場合にも、射の存在が保証される。

補題 1.8 (単調性条件をみたさない場合の射)**.** (V, T, μ) が条件 A をみたさないならば、(V, T, μ) よりも次元の小さいある対象 (V', T', μ') への射 $(V, T, \mu) \to (V', T', \mu')$ が存在する。

証明 末尾の付録に載せた。 □

補題 1.9 (条件 B をみたさない場合の射)**.** (V, T, μ) が条件 B をみたさないならば、(V, T, μ) よりも次元の小さいある対象 (V', T', μ') への射 $(V, T, \mu) \to (V', T', \mu')$ が存在する。

証明 末尾の付録に載せた。 □

以上の補題を用いて最小次元実現の特徴づけが得られる。

定理 1.10 (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) は条件 A と条件 B をみたす。

証明 (1) ⇒(2) 最小次元実現 (V, T, μ) が条件 A, B のいずれかをみたさなかったとすると、補題 1.8, 1.9 よりとくに (V, T, μ) よりも次元の小さい実現が存在することになり、矛盾が従う。

(2) ⇒(1) (V, T, μ) が条件 A と条件 B をみたすとする。$\mathcal{P}$ の任意の実現 (V', T', μ') に対し、補題 1.7 より全射線型写像 $L : V' \to V$ が存在するから、$\dim V \leq \dim V'$ である。したがって V は $\mathcal{P}$ の最小次元実現である。 □

例 1.11 (正規分布族の最小次元実現)**.** 定理 1.10 により、`0425_資料.pdf` の例 3.2 でみた正規分布族の例は最小次元実現であることがわかる。実際、$T(x) = {}^t(x, x^2)$ の像は $\mathbb{R}^2$ のいかなる真アファイン部分空間にも含まれないから、条件 A (3) が成り立つ。また、$\Theta^{\mathcal{P}} = \mathbb{R} \times \mathbb{R}_{<0}$ となることから条件 B も成り立つ。

本節の目標の定理を示す。

定理 1.12 (最小次元実現の間のアファイン変換)**.** $(V, T, \mu), (V', T', \mu')$ がともに最小次元実現ならば、(V, T, μ) から (V', T', μ') への射 (L, b) がただひとつ存在する。さらに、L は線型同型写像である。

証明 補題 1.4, 1.7 より、射 $(L, b)\colon (V, T, \mu) \to (V', T', \mu')$ はただひとつ存在する。また、補題 1.7 より存在する射 $(V', T', \mu') \to (V, T, \mu)$ をひとつ選んで (K, c) とおくと、合成射 $(K, c) \circ (L, b)$, $(L, b) \circ (K, c)$ は補題 1.4 より恒等射 $(\mathrm{id}_V, 0)$, $(\mathrm{id}_{V'}, 0)$ に一致する。したがって L は線型同型写像である。 □

同じことを圏の言葉を使わずに言い換えると次のようになる。

定理 1.13 (最小次元実現の間のアファイン変換). $(V,T,\mu),(V',T',\mu')$ がともに最小次元実現ならば、全射線型写像 $L:V\to V'$ とベクトル $b\in V'$ であって

$$T(x)=L(T'(x))+b \qquad \mu\text{-a.e.}x \tag{1.13}$$

をみたすものがただひとつ存在する。さらに、L は線型同型写像である。 □

系 1.14 (自然パラメータの変換). 上の定理の L は

$$\theta'(p)={}^tL(\theta(p)) \qquad (\forall p\in\mathcal{P}) \tag{1.14}$$

をみたす。ただし写像 $\theta\colon\mathcal{P}\to\Theta^{\mathcal{P}}$ および $\theta'\colon\mathcal{P}\to\Theta'^{\mathcal{P}}$ は P,P' (の $\Theta^{\mathcal{P}},\Theta'^{\mathcal{P}}$ 上への制限) の逆写像である。

証明 $p\in\mathcal{P}$ を任意とすると、指数型分布族の定義より、μ'-a.e.x に対し

$$\exp(\langle\theta(p),T(x)\rangle-\psi(\theta(p)))=\frac{dp}{d\mu}(x)=\exp(\langle\theta'(p),T'(x)\rangle-\psi'(\theta'(p))) \tag{1.15}$$

$$\therefore\quad \langle\theta(p),T(x)\rangle-\langle\theta'(p),T'(x)\rangle=\psi(\theta(p))-\psi'(\theta'(p)) \tag{1.16}$$

$$\therefore\quad \langle\theta(p),L(T'(x))+b\rangle-\langle\theta'(p),T'(x)\rangle=\psi(\theta(p))-\psi'(\theta'(p)) \tag{1.17}$$

$$\therefore\quad \langle{}^tL(\theta(p)),T'(x)\rangle-\langle\theta'(p),T'(x)\rangle=\psi(\theta(p))-\psi'(\theta'(p))-\langle\theta(p),b\rangle \tag{1.18}$$

$$\therefore\quad \langle{}^tL(\theta(p))-\theta'(p),T'(x)\rangle=\psi(\theta(p))-\psi'(\theta'(p))-\langle\theta(p),b\rangle \tag{1.19}$$

が成り立つ。したがって、(V',T',μ') の条件 A (2) より ${}^tL(\theta(p))=\theta'(p)$ が成り立つ。 □

今後の予定

- 指数型分布族 $\mathcal{P}$ 自体に構造を入れる。
- Amari-Chentsov テンソルを定義する。
- 正規分布族の場合の具体的な計算を行う (Fisher 計量、Levi-Civita 接続、測地線など)。

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.

[Yos] Taro Yoshino, **bn1970.pdf**, Dropbox.

付録

補題 1.8 の証明　(V,T,μ) が条件 A をみたさないという仮定から、ある $\theta \in V^\vee$, $\theta \neq 0$ および $r \in \mathbb{R}$ が存在して

$$\langle \theta, T(x) \rangle = r \qquad \mu\text{-a.e.}x \tag{1.20}$$

が成り立つ。そこで $V' := (\mathbb{R}\theta)^\perp = \{v \in V \mid \langle \theta, v \rangle = 0\}$ とおくと、ある可測写像 $T' \colon \mathcal{X} \to V'$ および $v_0 \in V$ が存在して $T(x) = T'(x) + v_0$ (μ-a.e.x) が成り立つ。このように定めた組 (V', T', μ) が $\mathcal{P}$ の実現であることは一旦認めて最後に示すこととし、まず次元と射について確かめる。

まず (V', T', μ) の次元は $\dim V' = \dim V - 1 < \dim V$ より (V,T,μ) の次元よりも小さい。また、射影 $\pi \colon V \to V'$ をひとつ選べば、$(\pi, 0)$ は明らかに (V,T,μ) から (V',T',μ) への射を与える。

あとは (V',T',μ) が $\mathcal{P}$ の実現であることを示せばよい。指数型分布族の定義 (0502_資料.pdf) の条件 (E0), (E1), (E2) は明らかに成立しているから、あとは条件 (E3) を確認すればよい。そこで $p \in \mathcal{P}$ を任意とする。いま (V,T,μ) が $\mathcal{P}$ の実現であることから、ある $\theta \in V^\vee$ が存在して

$$\frac{dp}{d\mu}(x) = \frac{\exp\langle \theta, T(x) \rangle}{\int_{\mathcal{X}} \exp\langle \theta, T(y) \rangle \, \mu(dy)} \qquad \mu\text{-a.e.}x \tag{1.21}$$

が成り立つ。T', v_0 を用いて式変形すると、μ-a.e.x に対し

$$\begin{aligned}
\frac{dp}{d\mu}(x) &= \frac{\exp(\langle \theta, T(x) \rangle)}{\int_{\mathcal{X}} \exp(\langle \theta, T(x) \rangle) \, \mu(dy)} &(1.22)\\
&= \frac{\exp(\langle \theta, T'(x) \rangle + \langle \theta, v_0 \rangle)}{\int_{\mathcal{X}} \exp(\langle \theta, T'(x) \rangle + \langle \theta, v_0 \rangle) \, \mu(dy)} &(1.23)\\
&= \frac{\exp(\langle \theta, T'(x) \rangle)}{\int_{\mathcal{X}} \exp(\langle \theta, T'(x) \rangle) \, \mu(dy)} &(1.24)
\end{aligned}$$

が成り立つ。したがって (V',T',μ) は条件 (E3) も満たし、$\mathcal{P}$ の実現であることがいえた。 □

補題 1.9 の証明　(V,T,μ) が条件 B をみたさないとする。すると、ある真ベクトル部分空間 $W \subsetneq V^\vee$ および $\theta_0 \in \Theta^{\mathcal{P}}$ が存在して $\operatorname{aspan}\Theta^{\mathcal{P}} = W + \theta_0$ が成り立つ。そこで $\widetilde{V} := V/W^\perp$ と定め、$\pi \colon V \to \widetilde{V}$ を自然な射影として $\widetilde{T} := \pi \circ T \colon \mathcal{X} \to \widetilde{V}$ と定める。また、$\mathcal{X}$ 上の測度 $\widetilde{\mu}$ を $\widetilde{\mu} := \exp\langle \theta_0, T(x) \rangle \cdot \mu$ と定める。このように定めた組 $(\widetilde{V}, \widetilde{T}, \widetilde{\mu})$ が $\mathcal{P}$ の実現であることは一旦認めて最後に示すこととし、まず次元と射について確かめる。

まず $(\widetilde{V}, \widetilde{T}, \widetilde{\mu})$ の次元は $\dim \widetilde{V} = \dim V - \dim W^\perp = \dim W < \dim V^\vee = \dim V$ より (V,T,μ) の次元よりも小さい。また、$(\pi, 0)$ は明らかに (V,T,μ) から $(\widetilde{V}, \widetilde{T}, \widetilde{\mu})$ への射を与える。

あとは $(\widetilde{V}, \widetilde{T}, \widetilde{\mu})$ が $\mathcal{P}$ の実現であることを示せばよい。指数型分布族の定義 (0502_資料.pdf) の条件 (E0), (E1), (E3) の成立は簡単に確かめられるから、ここでは条件 (E3) だけ確かめる。そこで $p \in \mathcal{P}$ を任意とする。(V,T,μ) が $\mathcal{P}$ の実現であることから、ある $\theta \in V^\vee$ が存在して

$$p(dx) = \frac{\exp\langle \theta, T(x) \rangle}{\int_{\mathcal{X}} \exp\langle \theta, T(x) \rangle \, d\mu(x)} \mu(dx) \tag{1.25}$$

が成り立つ。ここで線型写像 $\langle \theta - \theta_0, \cdot \rangle : V \to \mathbb{R}$ は $\mathrm{Ker}\,\langle \theta_0, \cdot \rangle \supset W^{\perp}$ をみたすから、図式

$$
\begin{tikzcd}
V \arrow[r, "\langle \theta - \theta_0, \cdot \rangle"] \arrow[d, two heads] & \mathbb{R} \\
\widetilde{V} \arrow[ru, dashed, "\widetilde{\theta}"'] &
\end{tikzcd}
\tag{1.26}
$$

を可換にする線型写像 $\widetilde{\theta} \colon \widetilde{V} \to \mathbb{R}$ すなわち線型形式 $\widetilde{\theta} \in \widetilde{V}^{\vee}$ が存在する。この $\widetilde{\theta}$ が条件 (E3) をみたすものであることを確かめればよいが、各 $x \in \mathcal{X}$ に対し

$$
\begin{aligned}
\left\langle \widetilde{\theta}, \widetilde{T}(x) \right\rangle &= \langle \theta - \theta_0, T(x) \rangle && (1.27) \\
&= \langle \theta, T(x) \rangle - \langle \theta_0, T(x) \rangle && (1.28)
\end{aligned}
$$

が成り立つから

$$
\begin{aligned}
p(dx) &= \frac{\exp \langle \theta, T(x) \rangle}{\int_{\mathcal{X}} \exp \langle \theta, T(x) \rangle \ \mu(dx)} \mu(dx) && (1.29) \\
&= \frac{\exp \left\langle \widetilde{\theta}, \widetilde{T}(x) \right\rangle \exp \langle \theta_0, T(x) \rangle}{\int_{\mathcal{X}} \exp \left\langle \widetilde{\theta}, \widetilde{T}(x) \right\rangle \exp \langle \theta_0, T(x) \rangle \ \mu(dx)} \mu(dx) && (1.30) \\
&= \frac{\exp \left\langle \widetilde{\theta}, \widetilde{T}(x) \right\rangle}{\int_{\mathcal{X}} \exp \left\langle \widetilde{\theta}, \widetilde{T}(x) \right\rangle \ \widetilde{\mu}(dx)} \widetilde{\mu}(dx) && (1.31)
\end{aligned}
$$

となる。したがって条件 (E3) の成立が確かめられた。以上より $(\widetilde{V}, \widetilde{T}, \widetilde{\mu})$ は $\mathcal{P}$ の実現である。これで証明が完了した。 □