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。

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

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。

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

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 1.12 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 $(L, b)(K, c) = (LK, Lc + b)$

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。

定義 1.2 (Affine span 条件). $\mathcal{P}$ 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 **affine span 条件**をみたすという。

命題-定義 1.3 (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) $P\colon V^\vee \to \mathcal{P}(\mathcal{X})$ 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 (1) $\iff$ (2) は `0502_資料.pdf` の命題 2.2 で示した。(2) $\iff$ (3) は `0523_板書.pdf` に書き留めてある。 □

単射性条件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) が単射性条件をみたすならば (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 V' の基底をひとつ選んで固定し、成分ごとに (V, T, μ) の単射性条件 (2) を適用すれば、$K = L$ を得る。よって (1.1) で $K = L$ として $b = c\ \mu$-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。

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 affine span 条件と単射性条件をみたすならば、任意の対象 (V', T', μ) から (V, T, μ) への射が存在する。

証明 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おく。

<u>Step 1: 射 (L, b) の構成</u> まず (V, T, μ) の affine span 条件より、$V^\vee$ のあるアファイン基底 $a^i \in \Theta^{\mathcal{P}}$ $(i =$

$0,\dots,m)$ が存在して、$e^i \coloneqq a^i - a^0$ $(i = 1,\dots,m)$ は $V^\vee$ の基底となる。そこで $p^i \coloneqq P(a^i) \in \mathcal{P}$ $(i = 0,\dots,m)$ とおき、(L, b) を次のように定める:

$$L\colon V' \to V, \quad t' \mapsto \left\langle \theta'(p^i) - \theta'(p^0), t' \right\rangle e_i \tag{1.3}$$

$$b \coloneqq \left\{ \psi(\theta(p^i)) - \psi(\theta(p^i)) - \psi(\theta'(p^0)) + \psi(\theta'(p^0)) \right\} e_i \in V \tag{1.4}$$

示すべきことは、すべての $p \in \mathcal{P}$ に対し

$$T(x) = L(T'(x)) + b \quad \mu\text{-a.e.}x \tag{1.5}$$

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

$$\exp(\langle \theta(p), T(x) \rangle - \psi(\theta(p))) = \frac{dp}{d\mu}(x) = \exp(\langle \theta'(p), T'(x) \rangle - \psi'(\theta'(p))) \qquad (x \in \mathcal{X} \setminus N_p) \tag{1.6}$$

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

が成り立つ。とくに各 $i = 0,\dots,m$ に対し

$$\left\langle \theta(p^i), T(x) \right\rangle - \left\langle \theta'(p^i), T'(x) \right\rangle = \psi(\theta(p^i)) - \psi'(\theta'(p^i)) \qquad (x \in \mathcal{X} \setminus N_{p^i}) \tag{1.8}$$

が成り立つから、

$$\begin{aligned} &\left\langle \theta(p^i) - \theta(p^0), T(x) \right\rangle - \left\langle \theta'(p^i) - \theta'(p^0), T'(x) \right\rangle \\ &= \psi(\theta(p^i)) - \psi'(\theta'(p^i)) - \psi(\theta(p^0)) + \psi'(\theta'(p^0)) \qquad (x \in \mathcal{X} \setminus (N_{p^0} \cup N_{p^i})) \end{aligned} \tag{1.9}$$

となる。ここで (V, T, μ) の単射性条件より $\theta(p^i) = \theta(P(a^i)) = a^i$ が成り立つから、上の式より

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

したがって

$$T(x) = L(T'(x)) + b \qquad (x \in \mathcal{X} \setminus (N_{p^0} \cup \cdots \cup N_{p^m})) \tag{1.11}$$

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) の単射性条件に反する。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 affine span 条件と単射性条件をみたすならば、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。 □

補題 1.8. (V,T,μ) が affine span 条件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 $(\widetilde{V},\widetilde{T},\widetilde{\mu})$ へのある射 $(V,T,\mu)\to(\widetilde{V},\widetilde{T},\widetilde{\mu})$ が存在する。

証明 (V,T,μ) が affine span 条件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

$$p(dx)=\frac{\exp\langle\theta,T(x)\rangle}{\int_{\mathcal{X}}\exp\langle\theta,T(x)\rangle\,d\mu(x)}\mu(dx)\tag{1.12}$$

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

$$\begin{tikzcd} V \arrow[r, "\langle\theta-\theta_0,\cdot\rangle"] \arrow[d, two heads] & \mathbb{R} \\ \widetilde{V} \arrow[ur, dashed, "\widetilde{\theta}"'] & \end{tikzcd}\tag{1.13}$$

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

$$\begin{aligned}\left\langle\widetilde{\theta},\widetilde{T}(x)\right\rangle&=\langle\theta-\theta_0,T(x)\rangle &(1.14)\\ &=\langle\theta,T(x)\rangle-\langle\theta_0,T(x)\rangle &(1.15)\end{aligned}$$

が成り立つから

$$\begin{aligned}p(dx)&=\frac{\exp\langle\theta,T(x)\rangle}{\int_{\mathcal{X}}\exp\langle\theta,T(x)\rangle\,\mu(dx)}\mu(dx) &(1.16)\\ &=\frac{\exp\left\langle\widetilde{\theta},\widetilde{T}(x)\right\rangle\exp\langle\theta_0,T(x)\rangle}{\int_{\mathcal{X}}\exp\left\langle\widetilde{\theta},\widetilde{T}(x)\right\rangle\exp\langle\theta_0,T(x)\rangle\,\mu(dx)}\mu(dx) &(1.17)\\ &=\frac{\exp\left\langle\widetilde{\theta},\widetilde{T}(x)\right\rangle}{\int_{\mathcal{X}}\exp\left\langle\widetilde{\theta},\widetilde{T}(x)\right\rangle\,\widetilde{\mu}(dx)}\widetilde{\mu}(dx) &(1.18)\end{aligned}$$

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。 □

補題 1.9. (V,T,μ) が単射性条件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 $(\widetilde{V},\widetilde{T},\widetilde{\mu})$ へのある射 $(V,T,\mu)\to(\widetilde{V},\widetilde{T},\widetilde{\mu})$ が存在する。

証明 (V, T, μ) よりも次元の小さい実現 $(\widetilde{V}, \widetilde{T}, \widetilde{\mu})$ の具体的な構成は、最小次元実現が単射性条件をみたすことを述べた `0425_コメント.pdf` の定理 0.3 (あるいは `0425_資料.pdf` の定理 4.2) の証明において与えた。その証明内の claim より射の構成も明らかである。 □

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 affine span 条件と単射性条件をみたす。

証明 (1) ⇒(2) 上の 2 つの補題と背理法により従う。

(2) ⇒(1) (V, T, μ) が affine span 条件と単射性条件をみたすとする。$\mathcal{P}$ 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。 □

注意 1.11 (正規分布族の最小次元実現). `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。単射性条件は、(θ_1, θ_2) が異なれば平均か分散の少なくとも一方が異なることになり、異なる確率分布を与えることからわかる。Affine span 条件も明らか。

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, b) は同型射であり、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 1.12 (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、線型写像 $L : V \to V'$ とベクトル $b \in V'$ であって

$$T(x) = L(T'(x)) + b \qquad \mu\text{-a.e.}x \tag{1.19}$$

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。

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。(L, b) の逆射は補題 1.7 より存在する射 $(K, c) \colon (V', T', \mu') \to (V, T, \mu)$ である。実際、合成射 $(K, c) \circ (L, b)$, $(L, b) \circ (K, c)$ は補題 1.4 より恒等射に一致する。 □

系 1.13 (自然パラメータの変換). 上の定理の L は

$$\theta'(p) = {}^t L(\theta(p)) \qquad (\forall p \in \mathcal{P}) \tag{1.20}$$

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' の逆写像である。

証明 $p \in \mathcal{P}$ を任意とすると、式 (1.7) のあたりと同様の議論により、μ'-a.e.x に対し

$$\langle \theta(p), L(T'(x)) + b \rangle - \langle \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) \tag{1.21}$$

$$\therefore \quad \langle {}^tL(\theta(p)), T'(x) \rangle - \langle \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) - \langle \theta(p), b \rangle \tag{1.22}$$

$$\therefore \quad \langle {}^tL(\theta(p)) - \theta'(p), T'(x) \rangle = \psi(\theta(p)) - \psi'(\theta'(p)) - \langle \theta(p), b \rangle \tag{1.23}$$

が成り立つ。したがって、(V', T', μ) の単射性条件より ${}^tL(\theta(p)) = \theta'(p)$ 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).

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