

1. 自己紹介

… 接触構造の極限として現れる葉層のトポロジーを知りたい人.

多様体 M の葉層 (foliation) $\mathcal{F}$ とは, M を葉と呼ばれる部分多様体 L_λ の非交和として表す分解 $M = \bigsqcup_{\lambda \in \Lambda} L_\lambda$ であって, M の各点の近くでは自明なファイブレーションであるものである. 各葉を点とする商空間 $M/\mathcal{F}$ (つまりパラメタ空間 Λ) を葉空間と呼ぶ. 葉空間が多様体になる場合がファイバー束である. 一般に可微分葉層の葉空間は商位相で Hausdorff 空間ですらないが, 上の局所ファイブレーションを集めたものを $M/\mathcal{F}$ の微分構造とすることができる. この意味で $M/\mathcal{F}$ に更に幾何構造が入る葉層が (ホロノミー不変な) 横断的構造を持つ葉層である. よく研究されているのは超曲面を葉とする余次元 1 葉層とベクトル場の流線による 1 次元葉層である. 基本的には力学系の幾何だから横断的幾何構造を考えるのは自然である.

M 上の非特異 1-形式 α が定める超平面場 $\xi = \ker \alpha$ が余次元 1 の葉層の接束 $T\mathcal{F}$ となる条件 (積分可能条件) は $\alpha \wedge d\alpha = 0$ である. 他方 M の次元を $2n+1$ として ξ が正の接触構造を定める条件は $\alpha \wedge (d\alpha)^n > 0$ (M の正の向き) である. 接触構造は葉層の対極にあるものであるが, 両者のトポロジーは不思議なほどよく似ている. とくに $n=1$ の場合には「 ≥ 0 」として葉層が接触構造の極限であるかのように考えるコンフォリエーションという枠組みがある. 接触構造は葉層よりも幾何的に複雑なものであるが, トポロジー的には Gray 安定性というものがあるので分かりやすい. Gray 安定性は接触構造のホモトピーがその容器である多様体のイソトピーによって追跡できることを主張するものである. 従って接触構造そのものを「結び目」ということができる. 超平面場のトポロジーの一部として見た場合, 葉層も接触構造も存在問題の答えはホモトピー論的であるが, 「結び目」は多様である. 葉層ではその難しさがよく分かる. というのも Reeb 安定性というものがあるが, 球体の境界の葉層が内部に伸びるためには, いかに内部に伸びる形をしていなければならない, そのように伸びるしかないからである. この現象に関連して, とくに 3 次元の場合には特別なトーラス葉の存在・非存在の区別が鍵となって葉層のトポロジーが展開した. ところが 4 次元以上の場合には与えられた葉層を変形して各葉が稠密であるような葉層をつくることのできる (最近の Meigniez の定理). 3 次元では存在しない場合も含めて特別なトーラス葉と接触構造のトポロジーを結びつけるコンフォリエーション理論があるので, それを高次元化すること (たとえば葉をシンプレクティック多様体とすること) は葉層のトポロジーを高次元化することにつながると考える. この研究は三松佳彦氏 (中央大) らによる Anosov 流の研究とシンプレクティック葉層 (Poisson 構造) の研究に関連して進行している.

参考文献

コンフォリエーションについては

本家 Y. Eliashberg, W. Thurston, Confoliations, Univ. Lect. Ser. 13 AMS 1998.

H. Kodama, Y. Mitsumatsu, S. Miyoshi, A. Mori, On Thurston's inequality for a spinnable foliation, Contemp. Math. 498 (2009) 195–206.

J. Bowden, Contact structures, deformations and taut foliations, Geom. Top. 20 (2016) 697–746.

M., The Reeb foliation arises as a family of Legendrian submanifolds at the end of

a deformation of the standard S^3 in S^5 , C. R. Acad. Sci. Paris 350 (2012) 67–70.
 葉層や接触構造の存在については
 G. Meigniez, Regularization and minimization of codimension-one Haefliger structures, J. Diff. Geom. 107 (2017) 157–202.
 M. Borman, Eliashberg, E. Murphy, Existence and classification of overtwisted contact structures in all dimensions, Acta Math. 215 (2015) 281–361.
 Anosov 流と非 Stein シンプレクティック多様体については
 Mitsumatsu, Anosov flows and non-Stein symplectic manifolds, Ann. Inst. Fourier 45 (1995) 1407–1421. (他 E. Ghys, Ann. Inst. Fourier 42 (1992), 209–247, H. Geiges, Int. Math. Res. Notices 1994, 23–30.
 P. Massot, K. Niederkrüger, C. Wendl, Weak and strong fillability of higher dimensional contact manifolds, Invent. Math. 192 (2013) 287–373.

2. 情報幾何とモジュライ理論

… 情報幾何とモジュライ理論の間には何か隠れた関係があるのかもしれない。

情報幾何はパラメトリック統計で扱うような確率分布の族についてのパラメタ空間の局所微分幾何である。正規分布の場合はその空間は上半平面 $\mathbb{H} = \{(m, s) \in \mathbb{R}^2 \mid s > 0\}$ (m は平均, s は標準偏差) である。そこに微分幾何的構造を定めるためにある種の「距離」を表す相対エントロピー $D : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}_{\geq 0}$ という関数を考え、それを対角線集合に制限したものの「3-jet」を見る。情報幾何が局所微分幾何であるのは、相対エントロピーの対角線集合のところだけを見たからだ(せいだ)と言ってしまおう。そうしてみると来るべき「大域情報幾何」では、先ず正規分布の対のパラメタ空間 $\mathbb{H} \times \mathbb{H}$ と関数 D のトポロジーを問題にすべきように思えてくる。

$\mathbb{H} \times \mathbb{H}$ をパラメタ空間と考えることは、数学では昔からやってきたことである。Abel 曲面は複素射影空間に埋め込むことができる複素 2 次元トーラスである。Abel 曲面の次の意味でのモジュライ空間は Hilbert モジュラー曲面と呼ばれる複素 2 次元多様体であり、それは $\mathbb{H} \times \mathbb{H}$ を Hilbert モジュラー群という不連続群で割ったものである：整数係数の多項式環 $\mathbb{Z}[x]$ を $x^2 + bx + c = 0$ という関係で割って得られる整環 (order) は判別式を $\Delta = b^2 - 4c$ として、 $\mathbb{R}$ の一部 $\mathbb{Q} + \mathbb{Q}\sqrt{\Delta}$ に自然に埋め込まれている。この整環を実乗法として持つ Abel 曲面は、幾何的には Δ から定まるあるシンプレクティック構造を保つように $\mathbb{C}^2$ を割った商である。その割り方は $\mathbb{H} \times \mathbb{H}$ の点を指定すれば定まり、商の Abel 曲面には誘導されたシンプレクティック構造が主偏極を定める。こうして得られる「実乗法を持つ主偏極アーベル多様体」の同型類を考えると、それは $\mathbb{H} \times \mathbb{H}$ の点を指定すれば定まり、 $\mathbb{H} \times \mathbb{H}$ のある商空間の点だと思える。この商を具体的に与えるのが Δ から定まる Hilbert モジュラー群というものであり、これに関する数論的・代数幾何的研究は伝統的で大きな蓄積を持っている。

… モジュライの「原理主義」というものがあって、「研究すべき」空間は別の「研究すべき」空間のモジュライ空間であるべきだとすれば、情報幾何を「研究すべき」理由として、正規分布の対と Abel 多様体には隠れた関係があるのかもしれない。

参考文献

次の論文に引用してある文献を見るとよい。

C. McMullen, Foliations of Hilbert modular surfaces, Amer. J. Math. 129(2007) 183–215.

3. モジュライ理論に関わる特異点論と接触トポロジー

… トポロジーは大域的なものを統一された全体として捉えることではなく、また大域的なあり方を忘れてただ局所的に捉えることでもない。それは対象をばらばらにしたり組み立てたりすることであり、そうやって何かを数えることである。

Hilbert モジュラー曲面のトポロジーは Hirzebruch によって研究された。それは元々曲面に属している特異点の解消だけでなく、カusp特異点を追加することによるコンパクト化と、そのカusp特異点の解消からなる研究である。このカusp特異点は Sol^3 -多様体という 3 次元幾何多様体の構成と関わるものである。その関わりの接触トポロジー的側面は粕谷直彦氏（京産大）によってはっきり書かれた。ここでは直ちに必要な部分だけを説明する。

整数の 2 次正方行列で行列式が 1 のもの $A \in SL(2, \mathbb{Z})$ を取り、 $\text{tr} A > 2$ とする。 A は実 2 次元トーラス T^2 の微分同相だと思えることができるので、 $T^2 \times [0, 1]$ の境界を $A: T^2 \times \{1\} \rightarrow T^2 \times \{0\}$ で貼り合わせて 3 次元多様体 N をつくることができる。これが Sol^3 -多様体と呼ばれるものである。行列 A は $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ と $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ を両方使って適当に並べた積に共役であることが知られている。ただし $SL(2, \mathbb{R})$ の元として共役と言ってるのではないからトレースが同じでも共役とは限らない。そして A を共役類の中で取り換えても N は変わらない。 $\text{tr}(A^{-1}) > 2$ でもあるから恒等的に 2 でない 2 以上の整数循環列 $\mathbb{Z}/r\mathbb{Z} \ni k \mapsto b_k \in \mathbb{Z}_{\geq 2}$ により A^{-1} は次と共役である。

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{b_{k+r-1}-2} \cdots \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{b_k-2}.$$

(k を動かしてよいのはシフトしても共役類は変わらないからである。) ここで公式

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{b_k-2} \right) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} b_k & 1 \\ -1 & 0 \end{bmatrix}$$

を使うと、 A^{-1} は次の積に共役であることが分かる。

$$\begin{bmatrix} s_k & t_k \\ u_k & v_k \end{bmatrix} := \begin{bmatrix} b_{k+r-1} & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} b_{k+r-2} & 1 \\ -1 & 0 \end{bmatrix} \cdots \begin{bmatrix} b_k & 1 \\ -1 & 0 \end{bmatrix}.$$

我々は共役類を見ているので、行列から始めるのではなく循環列 $b_k \geq 2$ で恒等的に 2 ではないものから始めてもよかったことになる。ここで連分数

$$w_k := b_k - \frac{1}{b_{k+1} - \frac{1}{-b_{k+2} - \cdots}} \left(= b_k - \frac{1}{w_{k+1}} \right)$$

を考えると、それは次の性質を持つことが容易にわかる。

$$\begin{bmatrix} s_k & t_k \\ u_k & v_k \end{bmatrix}^{-1} \begin{bmatrix} -1 \\ w_k \end{bmatrix} = w_1 \cdots w_n \begin{bmatrix} -1 \\ w_k \end{bmatrix}.$$

このことはまた w_k が次の整数係数 2 次方程式の解であることを意味する。

$$t_k w_k^2 - (s_k - v_k) w_k - u_k = 0 \quad (\text{判別式} = (s_k + v_k)^2 - 4 > 0)$$

他の解を $\bar{w}_k$ と書く。 $c = \log(w_1 \cdots w_r)$ とおき、 $K > 0$ に対して

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ w_1 & -\bar{w}_1 \end{bmatrix}^{-1} \begin{bmatrix} K & 0 \\ 0 & K \end{bmatrix}$$

とおくと、 $\mathbb{H} \times \mathbb{H}$ から次の同一視による商 $M(x, y, c) = \mathbb{H} \times \mathbb{H} / \sim$ が得られる．

$$(m, s, M, S) \sim (e^c m, e^c s, e^{-c} M, e^{-c} S),$$

$$(m, s, M, S) \sim (m + x_1, s, M + x_2, S),$$

$$(m, s, M, S) \sim (m + y_1, s, M + y_2, S)$$

ここで (m, s, M, S) は $\mathbb{H} \times \mathbb{H}$ の座標である．商が多様体になるのは写像

$$(m, s, M, S) \mapsto (e^c m, e^c s, e^{-c} M, e^{-c} S)$$

が mM 平面で $x = (x_1, x_2)$ と $y = (y_1, y_2)$ が生成する格子を保つことによる．格子に関して見るとこの写像は A と共役であり、従って $M(x, y, c)$ は Sol^3 -多様体 N を底面とする柱 $N \times \mathbb{R}$ と微分同相である．高さ $h \in \mathbb{R}$ は $h = -\log(sS)$ とする． h の定義域は $\mathbb{H} \times \mathbb{H} \subset \mathbb{C} \times \mathbb{C}$ で、Hessian $\text{diag} \left(\frac{1}{4s^2}, \frac{1}{4S^2} \right)$ は正定値だから、 h は狭義多重劣調和関数である．ここで我々は複素構造 J_{std} を忘れて次の原始形式が指定された完全シンプレクティック構造 $d\lambda_+$ について考えることにする．

$$\lambda_+ = -J_{\text{std}}^* dh = \frac{dm}{s} + \frac{dM}{S} = \exp \left(\frac{h+t}{2} \right) dm + \exp \left(\frac{h-t}{2} \right) dM$$

ただし $t = \log \frac{S}{s}$ ．そうすると上の多重劣調和性は次の方程式から定まる Liouville ベクトル場 X を用いて $dh(X) > 0$ と表すことができる．

$$\iota_X d\lambda_+ = \lambda_+.$$

このとき X の流線による葉層には λ_+ が定めるホロノミー不変な横断的接触構造がある．(この構造は接分布のトポロジーの言葉では偶接触構造、葉層は Cauchy 特性直線場というものだが、残念なことにその言葉では 1-形式の拡大を表す多重劣調和性が記述できない．) 特に断面 $\{h=0\}$ には λ_+ の制限が接触構造を定める．

一般に向きを持つ $2n+1$ 次元多様体 N 上に 1-形式 α があって $\alpha \wedge (d\alpha)^n > 0$ が成り立つとき、それは接触構造 $\xi = \ker \alpha$ を定める．このとき柱 $\mathbb{R} \times N$ 上には $\mathbb{R}$ の座標を z として $d(e^z \alpha)$ がシンプレクティック構造を定める．これをシンプレクティック化と呼び、 $\mathbb{R} \times \{*\}$ をその $\mathbb{R}$ -ファイバーと呼ぶ． $\mathbb{R}$ -ファイバーは我々がシンプレクティック化を考える場合にシンプレクティック形式よりもむしろその原始形式 $e^z \alpha$ を考えていることを表す．これを $z=0$ の断面に制限すれば元の接触形式 α が得られる．接触構造を定める 1-形式を $e^f \alpha$ ($f \in C^\infty(N)$) に取り換えることは接触形式となるための条件を変えないので、接触構造には横断的な向きを定めるのみで接触形式は自由に選ぶことが多い．この場合でも接触多様体のシンプレクティック化を $\mathbb{R}$ -ファイバー付き、ただし 0-断面は指定しない、と定義することは可能である．つまり接触形式 $e^f \alpha$ は関数のグラフ $z=f$ をあらためて 0-断面とするものである．

さて我々の $\mathbb{H} \times \mathbb{H}$ は $z=h/2$ とすると $\{z=0\} = \{Ss=1\}$ のシンプレクティック化である．ここで上のベクトル場 X の流線が $\mathbb{R}$ -ファイバーである．シンプレクティック化としての構造は商 $M(x, y, c)$ に遺伝し、 $\{z=0\}$ は商では Sol^3 -多様体 $N(x, y, c)$ である．これに点 ∞ を補ったものが Hilbert モジュラーカスプである．関数 t は $N(x, y, c)$ 上では T^2 -束の射影であり、 T^2 -ファイバーは mM -平面上で x と y が生成する格子による商として得られるものである． Mm 平面の標準的な座標でモノドロミー写像は $\begin{bmatrix} e^c & 0 \\ 0 & e^{-c} \end{bmatrix}$ であるが、これを格子で見れば $\begin{bmatrix} s_1 & t_1 \\ u_1 & v_1 \end{bmatrix}^{-1} \in SL(2, \mathbb{Z})$ だったということになる．そこで我々は最終的には次のような世界観に至る．

先ず $TN(x, y, c)$ に次の枠 (e_1, e_2, e_3) を取る．

$$e_1 = e^{-t/2} \partial_m, \quad e_2 = e^{t/2} \partial_M, \quad e_3 = 2\partial_t.$$

これは Lie 環 sol^3 を特徴づける次の関係式を満たす.

$$[e_1, e_2] = 0, \quad [e_3, e_1] = -e_1, \quad [e_3, e_2] = e_2.$$

双対余枠 (e_1^*, e_2^*, e_3^*) の方で見れば次のようになる.

$$de_3^* = 0, \quad de_1^* = -e_1^* \wedge e_3, \quad de_2^* = e_2^* \wedge e_3,$$

ここで

$$e_1^* = e^{t/2} dm, \quad e_2^* = e^{-t/2} dM, \quad e_3^* = dt/2$$

である. このとき $N(x, y, c)$ は A に付随する Sol^3 -多様体の構造を持つと言われる. e_3 が生成する懸垂流が Anosov 流であるのは, それが e_1 を縮小し e_2 を拡大するからである. その結果 2 つの 1-形式 $\alpha_{\pm} = e_1^* \pm e_2^*$ が定義する平面場 $\xi_{\pm}$ は互いに逆方向にねじれ, 正と負の接触構造となる. このうち正の接触構造 ξ_+ は上で定義したものと同一である. 負の接触構造のシンプレクティック化は

$$\lambda_- = e^{h/2}(e_1^* - e_2^*) = \exp\left(\frac{h+t}{2}\right) dm - \exp\left(\frac{h-t}{2}\right) dM.$$

の外微分である. これらの接触構造を Anosov 流で流していくと最終的には $\ker e_2^*$ が表す不安定葉層に収束するので, 少なくとも 3 次元多様体上のコンフォリエーション理論の素材となる. (実はそれだけでなく, 正と負のシンプレクティック化を持つことから, 5 次元多様体上でのコンフォリエーション理論の素材にもなる.)

参考文献

N. Kasuya, The canonical contact structure on the link of a cusp singularity, Tokyo J. Math. 37(2014) 1–20.

4. 情報幾何と正負の接触構造

… 正負の接触構造の共存は複素幾何の Levi 平坦性に関わると同時に, 安定・不安定方向に分離する双曲力学系と関わるものである. 情報幾何は 2 つの分布の違いを見るのだから, 2 つの方向に分離する原理があって然るべきである.

パラメトリック統計では, 開集合 $U \subset \mathbb{R}^n$ と, $\mathbb{R} \times U$ ($X = (X^1, \dots, X^n)$) 上の正の関数 $p(x, X) > 0$ で正規化条件

$$\int_{-\infty}^{\infty} p(x, X) dx = 1 \quad (\forall X \in U)$$

を満たすものを考える. ここで確率密度 $p_X(x) = p(x, X)$ を持つ確率変数 x のパラメタが X である. 絶対エントロピー.

$$- \int_{-\infty}^{\infty} p_X(x) \log p_X(x) dx$$

といったものをいきなり考えると, これは容易に負になりえるので, むしろもうひとつ別の確率変数を考えて, 相対エントロピー

$$D(X, Y) = - \int_{-\infty}^{\infty} p_X(x) \log \frac{p_Y(x)}{p_X(x)} dx \left(\geq - \log \int_{-\infty}^{\infty} p_X(x) \frac{p_Y(x)}{p_X(x)} dx = 0 \right)$$

というものが $(X, Y) \in U \times U$ 上で定まっていると考えるのが基本的である.

そもそも統計学には確率論的側面と情報理論的側面があり, どういうわけかそれらが互いに絡んでいる, というところが数学的に面白いのだと思う. 確率論的側面は 2 つの確率変数の間に演算を定めること, たとえば確率密度の積に定数を乗じて再び確率密度にするとか, 確率密度の畳み込みを確率密度にするとかが最初の目的であ

り、中心極限定理は普通そうした代数的な興味から出てくるものだと考えられている。これに対して情報理論的側面は2つの確率変数の間に「距離」を定めることが最初の目的であり、甘利氏の情報幾何はまさにそうした幾何的な興味から出てきたものだと思う。甘利氏の情報幾何では $U \times U$ の対角線集合に沿って、Fisher 情報量（普通は Riemann 計量であるもの） $g = \sum_{i,j} g_{ij} dX^i dX^j$ が2次近似

$$D(X, X + \Delta X) \approx \frac{1}{2!} \sum_{i,j} g_{ij} \Delta X^i \Delta X^j (\approx D(X + \Delta X, X))$$

から出てきて、更に（江口氏の解釈では言うべきらしいが）e-接続 ∇ が3次近似

$$\begin{aligned} & D(X + \Delta X, X) - D(X, X + \Delta X) \\ & \approx \frac{1}{3!} \sum_{i,j,k} (\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} - 2\Gamma_{ij,k}) \Delta X^i \Delta X^j \Delta X^k \end{aligned}$$

から出てくる。ここで

$$\begin{aligned} |\Delta X| &= \sqrt{(\Delta X^1)^2 + \cdots + (\Delta X^n)^2} \ll 1, \\ \nabla_{\partial_i} \partial_j &= \sum_k \Gamma_{ij}^k \partial_k, \quad \Gamma_{ij,k} = \sum_l g_{kl} \Gamma_{ij}^l, \quad g_{ij} = g_{ji}, \quad \text{and} \\ \partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} - 2\Gamma_{ij,k} &= \partial_j g_{ki} + \partial_k g_{ij} - \partial_i g_{jk} - 2\Gamma_{jk,i}. \end{aligned}$$

である。つまり U 上のアフィン接続の空間のベクトルで Levi-Civita 接続を始点として e-接続を終点とするものが $D(X, Y)$ の非対称性を対角線集合に沿って1次近似したものとして定まるというのが e-接続の定義である。このベクトルを基本ベクトルとする数直線を取り、実数 α が表す接続を α -接続と呼ぶ。幾何学者・トポロジストとして付言すると単に式で係数をこう決めるという定義の仕方は止めるべきであり、定義から明らかなように接続の差がテンソルだということと、関数またはベクトルの最低次のジェットはテンソルだということ、計量を持っているなら Levi-Civita 接続も持っているということから上のようにして α -接続を導入すべきである。

さて正規分布の話に戻って、 (m, s, M, S) を $\mathbb{H} \times \mathbb{H}$ の座標とし、対応

$$\begin{aligned} F: \quad M &= -\frac{m}{s^2} \quad \text{and} \quad S = \frac{1}{s} \\ \left(\Leftrightarrow \quad \frac{m}{s} + \frac{M}{S} &= 0 \quad \text{and} \quad sS = 1 \right) \end{aligned}$$

を $\mathbb{H} \times \mathbb{H}$ の部分多様体とみなす。任意の点 $(m, s, M, S) \in \mathbb{H} \times \mathbb{H}$ に対して

$$f(m, s, M, S) = \frac{\left(\frac{M}{S} + \exp(-h) \frac{m}{s} \right)^2 + \exp(-2h) - 1 + 2h}{2},$$

とおく。ただし $h = -\log(sS)$ とした。このとき

命題 1. $\mathbb{H} \times \mathbb{H}$ の第1成分の点 (m', s') と第2成分の点 (M, S) を $(m', s', M, S) \in F$ で対応させるとき、 $f(m, s, M, S)$ は相対エントロピー $D((m, s), (m', s'))$ を表す。

証明.

$$\begin{aligned} 2f\left(m, s, \frac{-m'}{(s')^2}, \frac{1}{s'}\right) &= \left(\frac{-m'}{s'} + \frac{s}{s'} \cdot \frac{m}{s} \right)^2 + \frac{s^2}{(s')^2} - 1 - 2\log \frac{s}{s'} \\ &= 2D((m, s), (m', s')) \end{aligned}$$

□

積 $(m, s) * (m', s')$ を次のように定義すれば、確率密度の畳み込みを表す。

$$(m, s) * (m', s') = \left(m + m', \sqrt{s^2 + (s')^2} \right).$$

また積 $(M, S) \cdot (M', S')$ を次のように定義すれば、確率密度の積の正規化になる。

$$(M, S) \cdot (M', S') = \left(\frac{M(S')^2 + (M')S^2}{S^2 + (S')^2}, \sqrt{\frac{S^2(S')^2}{S^2 + (S')^2}} \right).$$

対応 F は次のようにこれらの演算を入れ替えるので Fourier 変換に似ている。

命題 2. $F(m, s, M, S) = F(m', s', M', S') = 0$ として $(m'', s'') = (m, s) * (m', s')$, $(M'', S'') = (M, S) \cdot (M', S')$ ならば $F(m'', s'', M'', S'') = 0$.

証明.

$$\begin{aligned} M'' &= \frac{M(S')^2 + (M')S^2}{S^2 + (S')^2} = \frac{-\frac{m}{s^2} \frac{1}{(s')^2} - \frac{m'}{(s')^2} \frac{1}{s^2}}{\frac{1}{s^2} + \frac{1}{(s')^2}} \\ &= \frac{-(m + m')}{s^2 + (s')^2} = -\frac{m''}{(s'')^2}, \\ S'' &= \sqrt{\frac{S^2(S')^2}{S^2 + (S')^2}} = \frac{1}{\sqrt{\frac{1}{s^2} + \frac{1}{(s')^2}}} = \frac{1}{s''}. \end{aligned}$$

□

この対応 F は正規分布の統計における確率論的側面ないし代数的側面を表すものであるが、他方で情報理論的側面、幾何的側面を表す相対エントロピーを対応 F のもとで書き換えた関数 f の形は、 F の形とともに接触幾何的な対称性を持つ。

命題 3. 対応 F と関数 f はモノドロミー写像 $(m, s, M, S) \mapsto (e^c m, e^c s, e^{-c} s, e^{-c} S)$ で不変である。

証明. モノドロミー写像は逆変動係数 $\frac{m}{s}, \frac{M}{S}$ と関数 $h = -\log(sS)$ を保つ。 □

命題 4. $\mathbb{H} \times \mathbb{H}(\ni (m, s, M, S))$ 上のシンプレクティック構造

$$d\lambda_+ = \frac{dm \wedge ds}{s^2} + \frac{dM \wedge dS}{S^2},$$

フィッシャー情報量の和

$$g = \frac{dm^2 + 2ds^2}{s^2} + \frac{dM^2 + 2dS^2}{S^2},$$

および概複素構造

$$J : \partial_m \mapsto \frac{1}{\sqrt{2}} \partial_s, \partial_M \mapsto \frac{1}{\sqrt{2}} \partial_S$$

は $g(\cdot, \cdot) = \sqrt{2} d\lambda_+(\cdot, J\cdot)$ を満たし、商 $M(x, y, c)$ に遺伝する。

証明. $\sqrt{2} d\lambda_+(\partial_m, \frac{1}{\sqrt{2}} \partial_s) = \frac{1}{s^2}, \sqrt{2} d\lambda_+(\partial_s, -\sqrt{2} \partial_m) = \frac{2}{s^2}$. 残りは容易。 □

命題 5. 制限 $\alpha_+ = \lambda_+|_{\tilde{N}}$ を接触形式として持つ $(\mathbb{H} \times \mathbb{H}, d\lambda_+)$ の超曲面 $\tilde{N} = \{h = 0\}$ の中で曲面 F は Legendre 直線

$$\{e^{T/2}m + e^{-T/2}M = 0 \quad \text{and} \quad t = T\} \quad (T \in \mathbb{R}).$$

の非交和（葉層）であり、この構造は商 $N(x, y, c)$ への F の稠密はめ込みに遺伝する。

証明. F の定義式は $h = 0$ と $e^{t/2}m + e^{-t/2}M = 0$ である. $\square$

命題 6. 負のシンプレクティック構造

$$d\lambda_- = \frac{dm \wedge ds}{s^2} - \frac{dM \wedge dS}{S^2}$$

に関して F は Lagrange 部分多様体である。

証明. $TF = \text{span}\{s\partial_m - S\partial_M, m\partial_m + s\partial_s - M\partial_M - S\partial_S\}$ であり,

$$\frac{dm \wedge ds(s\partial_m, m\partial_m + s\partial_s)}{s^2} - \frac{dM \wedge dS(-S\partial_M, -M\partial_M - S\partial_S)}{S^2} = 1 - 1 = 0.$$

$\square$

以上は接触幾何的性質だが、次の情報幾何的性質も興味深い。

命題 7. F は $\mathbb{H} \times \mathbb{H}$ の第 1 成分の e-測地線（e-接続の測地線）と第 2 成分 $\mathbb{H}$ の e-測地線を（パラメタ付き曲線としてではなく像としてではあるが）対応させる。

証明. ここでは ms -半平面の縦座標を $v = s^2$ に変える. このとき $\left(-\frac{m}{v}, \frac{1}{v}\right)$ はアフィン座標であるから、測地線の像は直線（と上半平面の共通部分）であり、そのアフィンパラメタは $1/v$ である. アフィン座標を F でうつすと (M, V) となるから ($V = S^2$), 測地線は標準的なアフィンパラメタを持つ直線と SV -半平面の共通部分に写される. これは SV -半平面の e-測地線と像として同じものである. $\square$

参考文献

本節と次節の内容は次のプレプリントに書いたことです（Ocami 版は古いです）.
M., Information geometry in a global setting, preprint.

5. 接触 HAMILTON 流と BAYES 推定

… 統計学の推論の正しさを流線の滑らかさという観点から評価する。

Bayes 統計においては観測は事象であり、母数 x は確率変数である。母分布は未知だから、 x の従う分布も未知であり、それは事前分布として決めるべきものである。 x の事前分布の密度を $p(x)$ とする。観測は事象なので確率を持つ。それを r とする。観測は各 x に対応する母分布を決めた場合の条件つき確率を持つので、それを x の関数とみなして尤度と言う。尤度は単に非負の関数であるが、定数で正規化することによって確率密度 $q(x)$ が得られたとする。このとき観測を条件とする条件付確率（事後確率）の密度 $p'(x) = p(x)q(x)/r$ が Bayes の法則から得られる。観測をする前の「思い込み」である事前分布を観測をもとに事後分布に修正することを Bayes 学習と呼ぶ。Bayes 学習を繰り返して得られた事後分布を用いて現実の判断を行うことを Bayes 推定と呼ぶ。分散を既知としたときの平均の Bayes 推定に用いるのは、 $p(x)$ と $q(x)$ が正規分布であるような学習である。このとき $p'(x)$ を得る演算は積「 $\cdot$ 」である。「 $\cdot$ 」の冪は同じ分布を学習し続けることを表す。この過程に F で対応するのが、畳み込みから来る積「 $*$ 」の冪である。それは mv 平面上で原点を端とする直線上を一定の間隔で進む。これを F で写すと MV 平面上で $M = \text{一定の線}$ に沿って M 軸

上の点に収束する過程が得られる。こちらが Bayes 学習の過程であり、推定される平均 M に確率密度が集中していく様子として解釈することができる。本節ではこの過程をある接触 Hamilton ベクトル場が生成する流の離散化として記述する。

そもそも接触 Hamilton ベクトル場とは何かを説明する。多様体 N 上の接触形式 α と関数 H からシンプレクティック化 $(\mathbb{R} \times N, d(e^z H))$ と同次化 $\tilde{H} = e^z H$ をつくり、 $\tilde{H}$ を Hamilton 関数とする $\mathbb{R} \times N$ 上の Hamilton ベクトル場 $\tilde{Y}$ を取る。このとき容易に分かるように $[\partial_z, \tilde{Y}]$ は ∂_z と平行なので、 $\tilde{Y}$ を $\mathbb{R}$ -ファイブレーションによって N 上のベクトル場 Y に押し出すことができる。このベクトル場は 0-断面上にあると思っていので、 $\alpha(Y)$ を計算すると、元の関数 H が出てくる。このとき

$$\mathcal{L}_{\tilde{Y}} e^z \alpha = \iota_{\tilde{Y}} d(e^z \alpha) + d\iota_{\tilde{Y}}(e^z \alpha) = -d(e^z H) + d(e^z H) = 0$$

であるから Y は接触構造 $\xi = \ker \alpha$ を保つベクトル場つまり接触ベクトル場である。接触構造 ξ の非可積分性から、 ξ にどこでも接する接触ベクトル場は 0 のみであり、従って $\alpha(Y) = H$ を満たす接触ベクトル場 Y は一意に定まる。よって勝手な接触ベクトル場 Y は接触 Hamilton 関数 $\alpha(Y)$ に関する接触 Hamilton ベクトル場として解釈される。ここで $\tilde{H}$ を固定したまま、 $\mathbb{R}$ -ファイブレーションの 0-断面を取り換えることは新たに接触形式 $e^\varphi \alpha$ と接触 Hamilton 関数 $e^{-\varphi} H$ を考えることであるから、接触形式は接触ベクトル場全体から関数全体への線形同型写像を決めるものであり、その決め方には正の関数倍の自由度があることが分かる。(このことから接触構造を固定しない場合の接触ベクトル場は Hamilton 関数そのものではなく、その符号と関係することが分かる。それが Giroux の凸接触幾何の基本的なアイデアである。)

定理 1. 逆変動係数 $\frac{m}{s}$ を $\mathbb{H} \times \mathbb{H}$ 上の関数と見て接触型超曲面 $(\tilde{N}, \alpha_+)$ に制限したとき、それに関する接触 Hamilton ベクトル場 Y は次のように表される。

$$Y = \frac{m}{s} e_1 - \frac{1}{2} e_3 = m \partial_m - \frac{1}{2} (S \partial_S - s \partial_s)|_{\tilde{N}}.$$

このベクトル場は曲面 F に接するので、 F への制限 $Y|_F$ を考え、それを $\mathbb{H} \times \mathbb{H}$ の第 1 成分と第 2 成分に射影して得られるベクトル場を Y_1, Y_2 とする。このとき Y_1 の各流線は原点 $(0, 0) \in \mathbb{H}$ を端とする e-測地線 (の像) であり、 Y_2 の各流線は $M = \text{一定}$ の e-測地線 (の像) である。(よって時間を刻む間隔を調整するか、あるいは時間に依存して Hamilton 関数の大きさを変えることにより、上記の Bayes 学習の過程は Y が生成する接触 Hamilton 流の離散化であると言えることになる。)

証明. 次の内積 $\iota_Y d\left(\frac{m}{s} + \frac{M}{S}\right) = \frac{1}{2} \left(\frac{m}{s} + \frac{M}{S}\right)$ は F に沿って消える。ここで $\mathcal{L}_Y \alpha_+ = \frac{1}{2} \alpha_+$ であり、Hamilton 関数は $\iota_Y \alpha_+ = \frac{m}{s}$ である。□

6. STUDENT- t 分布への移行

位置 $m \in \mathbb{R}$, 大きさ $s \in \mathbb{R}_{>0}$, 自由度 $\nu \geq 1$ の Student- t 分布の密度は

$$p_\nu(m, s)(x) = \frac{1}{B\left(\frac{\nu}{2}, \frac{1}{2}\right) \sqrt{\nu s}} \left\{ 1 + \frac{(x - m)^2}{s^2 \nu} \right\}^{-\frac{\nu+1}{2}},$$

で与えられる (B はベータ関数) . このとき相対エントロピーは

$$\begin{aligned} D_\nu((m, s), (m', s')) &= \int_{-\infty}^{\infty} p_\nu(m, s) \log \frac{p_\nu(m, s)}{p_\nu(m', s')} dx \\ &= -\log \frac{s}{s'} + \frac{\nu+1}{2B(\frac{\nu}{2}, \frac{1}{2})} \int_{-\infty}^{\infty} \frac{1}{(1+x^2)^{\frac{\nu+1}{2}}} \log \frac{1 + \left(\frac{s}{s'}x + \frac{m-m'}{\sqrt{\nu s'}} \right)^2}{1+x^2} dx, \end{aligned}$$

であるから次ように関数 f_ν を取れば今までの話と同様のことができる.

$$\begin{aligned} f_\nu(m, s, M, S) &= -\log(sS) + \frac{\nu+1}{2B(\frac{\nu}{2}, \frac{1}{2})} \int_{-\infty}^{\infty} \frac{1}{(1+x^2)^{\frac{\nu+1}{2}}} \log \frac{1 + \left(sSx + \frac{sS\frac{m}{s} + \frac{M}{S}}{\sqrt{\nu}} \right)^2}{1+x^2} dx \end{aligned}$$

心配なのは計量や測地線であるが³, 以下のように考えると解決してしまう. Fisher 計量を g_{ij} , e-接続の係数を Γ_{jk}^i , Levi-Civita 接続の係数を $\{\overset{i}{jk}\}$ とすると

$$\begin{aligned} g_{11} &= \frac{1}{s^2} \cdot \frac{\nu+1}{\nu+3}, & g_{22} &= \frac{2}{s^2} \cdot \frac{\nu}{\nu+3}, & g_{12} &= g_{21} = 0, \\ \Gamma_{11}^1 &= 0, & \Gamma_{22}^1 &= 0, & \Gamma_{12}^1 &= \Gamma_{21}^1 = \frac{-2}{s} \cdot \frac{\nu+2}{\nu+5}, \\ \Gamma_{11}^2 &= \frac{3}{s\nu} \cdot \frac{\nu+1}{\nu+5}, & \Gamma_{22}^2 &= \frac{-3}{s} \cdot \frac{\nu+1}{\nu+5}, & \Gamma_{12}^2 &= \Gamma_{21}^2 = 0, \\ \{\overset{1}{11}\} &= 0, & \{\overset{1}{22}\} &= 0, & \{\overset{1}{12}\} &= \{\overset{1}{21}\} = \frac{-1}{s} \\ \{\overset{2}{11}\} &= \frac{1}{2s} \cdot \frac{\nu+1}{\nu}, & \{\overset{2}{22}\} &= \frac{-1}{s}, & \{\overset{2}{12}\} &= \{\overset{2}{21}\} = 0. \end{aligned}$$

ここで $\lim_{\nu \rightarrow \infty} g_{ij}$ と $\lim_{\nu \rightarrow \infty} \Gamma_{jk}^i$ は正規分布のときに出てきたものである.

定理 2. Student- t 分布の空間を $\mathbb{H}_\nu$ として, その上の α -接続 $\alpha\Gamma_{jk}^i + (1-\alpha)\{\overset{i}{jk}\}$ は $\nu \rightarrow \infty$ とすると正規分布の空間 $\mathbb{H}_\infty$ の α -接続に収束する. 1 より大きい任意の ν について, $\mathbb{H}_\nu$ 上の $\frac{\nu+5}{\nu-1}$ -接続は $\mathbb{H}_\infty$ 上の e-接続と一致する. つまりここが不動点となってペンシルのように α -接続のなす直線が $1 < \nu \leq \infty$ の範囲で回転する.

証明.

$$\begin{aligned} \frac{\nu+5}{\nu-1}\Gamma_{12}^1 + \frac{-6}{\nu-1}\{\overset{1}{12}\} &= \frac{-2(\nu+2)+6}{s(\nu-1)} = \frac{-2}{s} = \lim_{\nu \rightarrow \infty} \Gamma_{12}^1, \\ \frac{\nu+5}{\nu-1}\Gamma_{11}^2 + \frac{-6}{\nu-1}\{\overset{2}{11}\} &= \frac{3(\nu+1)-3(\nu+1)}{s\nu} = 0 = \lim_{\nu \rightarrow \infty} \Gamma_{11}^2, \\ \frac{\nu+5}{\nu-1}\Gamma_{22}^2 + \frac{-6}{\nu-1}\{\overset{2}{22}\} &= \frac{-3(\nu+1)+6}{s(\nu-1)} = \frac{-3}{s} = \lim_{\nu \rightarrow \infty} \Gamma_{22}^2. \end{aligned}$$

ここで 0-接続 $\{\overset{i}{jk}\}$ が正規分布 e-接続を通らない直線上を動くことに注意する. $\square$

このようなわけで正規分布の e-接続を $\mathbb{H}_\nu$ に共通の接続として取れば, 前節までの結果が Student- t 分布についても成り立つことになる. 細かいことを注意すると,

命題 8. Fisher 計量の和

$$G = \frac{(\nu+1)dm^2 + 2\nu ds^2}{(\nu+3)s^2} + \frac{(\nu+1)dM^2 + 2\nu dS^2}{(\nu+3)S^2},$$

は Riemann 計量であり, 概複素構造

$$J: \partial_m \mapsto \sqrt{\frac{\nu+1}{2\nu}} \partial_s, \quad \partial_M \mapsto \sqrt{\frac{\nu+1}{2\nu}} \partial_S$$

を取ると $G(\cdot, \cdot) = \frac{\sqrt{2\nu(\nu+1)}}{\nu+3} d\lambda_+(\cdot, J\cdot)$ が成り立つ.

証明.

$$\begin{aligned} \frac{\sqrt{2\nu(\nu+1)}}{\nu+3} d\lambda_+ \left(\partial_m, \sqrt{\frac{\nu+1}{2\nu}} \partial_s \right) &= \frac{\nu+1}{(\nu+3)s^2} \\ \frac{\sqrt{2\nu(\nu+1)}}{\nu+3} d\lambda_+ \left(\partial_s, -\sqrt{\frac{2\nu}{\nu+1}} \partial_m \right) &= \frac{2\nu}{(\nu+3)s^2}. \end{aligned}$$

□

学習についても Student- t バージョンがあり, 分散が未知の場合の平均の Bayes 推定に用いられる. x_n は $N(\mu_n, 1/\tau)$ に従う確率変数の列であるとして, 母平均 μ_n は「滑らか」に動き, 母分散 $1/\tau$ は一定であるとする. 母平均が一定である場合と同様, 事前分布としてガンマ正規分布 $\Gamma N\left(\alpha_n, \beta_n, m_n, \frac{1}{t_n}\right)$ を取り, データ x_n から

事後分布 $\Gamma N\left(\alpha'_n, \beta'_n, m'_n, \frac{1}{t'_n}\right)$ を得たとする

$$\begin{aligned} m'_n &= \frac{1}{t_n+1} x_n + \frac{t_n}{t_n+1} m_n, & t'_n &= t_n + 1, \\ \alpha'_n &= \alpha_n + \frac{1}{2}, & \beta'_n &= \beta_n + \frac{t_n}{t_n+1} \cdot \frac{(x_n - m_n)^2}{2}. \end{aligned}$$

である. このとき Student- t 周辺分布は $\mathbb{H}_{n-1}$ 上の点 $\left(m'_n, \sqrt{\frac{\beta'_n}{t'_n \alpha'_n}}\right)$ として表される. これを $\mathbb{H} \times \mathbb{H}$ の第 2 成分とする. 各ステップでの事前分布の決め方を

$$\begin{aligned} m_n &= \varepsilon_n, & t_n &= t_{n-1} \quad (t_0 = 0), \\ \alpha_n &= \alpha_{n-1} \quad (\alpha_0 = 0), & \text{and} \quad \beta_n &= \beta_{n-1} \quad (\beta_0 = 0), \end{aligned}$$

として, ε_n については F をそれだけずらした場合に対応する第 1 成分の点

$$\left(-\frac{\alpha'_n t'_n (m'_n - \varepsilon_n)}{\beta'_n}, \sqrt{\frac{\alpha'_n t'_n}{\beta'_n}} \right)$$

が (s -軸というよりむしろ) v -軸に近くなるようにアルゴリズムをつくる. たとえば

$$\varepsilon_n = 0.1x_{n-2} + 0.2x_{n-1} + 0.4x_n + 0.2x_{n+1} + 0.1x_{n+2} \quad (x_{-1} = x_0 = x_1)$$

とする. もし ε_n が滑らかに動くのであれば, 第 1 成分の点は v 軸の近くをおよそ単調な仕方で動くはずであるから, そこを見て推定が上手くいったかどうかを判定する.

参考文献 M., Symplectic geometry of Student-t distributions, 準備中.