計量の退化を許容する統計多様体

北海道大学　大学院情報科学学院　情報理工学コース
嘉陽海渡 (Kaito KAYO) *

概 要

統計多様体はコダッチ構造を持つリーマン多様体 (i.e., 計量が非退化) のことであり，さまざまな分野での応用が見られる．しかし，実応用上では Fisher 計量のように退化してしまうケースが存在し，その場合コダッチ構造が現れない．本研究は，退化計量を許容するように統計多様体の概念を一般化するものである．

1　導入

情報幾何とは特別な形をした統計多様体の理論のことであり，機械学習などの分野で応用されている．しかし，実応用上では計量が退化してしまう状況がしばしば現れる．例として，ディープラーニングに用いられるニューラルネットや混合ガウス分布においては計量が退化してしまう場合があることがわかっている．この状況においては統計多様体，または情報幾何の理論は適用できなくなってしまう．しかし，Nakajima-Ohmoto[8] によって，計量が退化する場合における情報幾何の拡張 (**概ヘッセ構造**) が導入された．本稿では，統計多様体の理論を計量が退化する場合に拡張 (**概コダッチ構造**) し，さらに特別な状況では概ヘッセ構造と一致していることを確認する．

2　統計多様体

本章では統計多様体の定義とその性質を解説する．より詳しい解説は [1, 2, 7] を参照せよ．(M, h) を n 次元擬リーマン多様体とし，∇ を M 上のアファイン接続とする．また,$X, Y, Z \in \Gamma(TM)$ とする．

定義 2.1 [1, 2, 7] ∇ の計量 h に関する双対接続 ∇^* を次で定義する．

$$Xh(Y, Z) = h(\nabla_X Y, Z) + h(Y, \nabla^*_X Z).$$

h の非退化性より ∇^* は一意に存在し，明らかに $(\nabla^*)^* = \nabla$ が成り立つ．

命題 2.2 (Matsuzoe [7, Proposition2.2]) 次の 4 つの条件の内，2 つを満たせば残りの 2 つも同時に満たされる．

(i) ∇ が torision-free (i.e., $\nabla_X Y - \nabla_Y X - [X, Y] = 0$);

(ii) ∇^* が torsion-free (i.e., $\nabla^*_X Y - \nabla^*_Y X - [X, Y] = 0$);

*E-mail:kayo.kaito.b9@elms.hokudai.ac.jp

(iii) $C := \nabla h$ が全対称 (i.e., $C(X,Y,Z) = C(Y,X,Z) = C(X,Z,Y)$);

(iv) $\nabla^h = (\nabla + \nabla^*)/2$ が成り立つ．ただし ∇^h は h に関する Levi-Civita 接続.

定義 2.3 [1, 2, 7] 命題 2.2 の 4 つの条件を全て満たすとき，すなわち，どれか 2 つの条件が成り立っているとき，(h, ∇, ∇^*) を**統計構造**，または**コダッチ構造**と呼び，統計構造を持つ M を**統計多様体**と呼ぶ．特に ∇ と ∇^* がどちらも平坦 (曲率が常に 0) の時には**双対平坦構造**，または**ヘッセ構造**と呼ばれる．

例 2.4 (h, ∇^h, ∇^h) は M 上の統計構造である．実際，Levi-Civita 接続は計量を保つので，

$$Xh(Y,Z) = h(\nabla^h_X Y, Z) + h(Y, \nabla^h_X Z)$$

を満たし，∇^h の双対接続は ∇^h になる．さらに，Levi-Civita 接続は torsion-free であるので，(h, ∇^h, ∇^h) は統計構造になっている．

例 2.5 指数型分布族

$$M = \left\{ p(x|\boldsymbol{\theta}) = \exp\left\{ C(x) + \sum_{i=1}^{n} \theta^i F_i(x) - \psi(x) \right\} \mid \boldsymbol{\theta} \in \mathbb{R}^n \right\}$$

を考える．ただし，$x \in \mathbb{R}$ は確率変数，C と各 F_i は $\mathbb{R}$ 上の滑らかな関数，ψ は規格化因子である．M はパラメータ $\boldsymbol{\theta}$ を n 次元の座標系と思うことで n 次元多様体になる．また，$(0,2)$ テンソル h を

$$h\left(\frac{\partial}{\partial\theta^i}, \frac{\partial}{\partial\theta^j}\right) = \int_{\mathbb{R}} \left(\frac{\partial}{\partial\theta^i} \log p(x|\boldsymbol{\theta})\right)\left(\frac{\partial}{\partial\theta^j} \log p(x|\boldsymbol{\theta})\right) p(x|\boldsymbol{\theta}) dx$$

で定めると，h は非退化かつ対称であり M 上の計量になっていることがわかる．この h は Fisher 計量と呼ばれる．アファイン接続 ∇ を，各 i, j で

$$\nabla_{\frac{\partial}{\partial\theta^i}} \frac{\partial}{\partial\theta^j} = 0$$

を満たすようにとると，∇ は平坦かつ torsion-free である．さらに，h に関する双対接続 ∇^* も平坦かつ torision-free であることが確かめることができる．したがって，(h, ∇, ∇^*) は M 上の双対平坦構造である．

3　概コダッチ構造

統計構造は計量の非退化性を用いて定義された．もし計量が退化しているとすると，双対接続が定義されないので統計構造を考えることができない．本章では計量の退化を許容するように統計構造の定義を拡張する．

定義 3.1 [6, Definition 3] M を n 次元多様体，E を M 上の $2n$ 階のベクトル束，τ を E 上の擬リーマン計量とし，$I : E \to E$ を束写像とする．3 つ組 (E, τ, I) が次の条件を満たす時，**パラエルミートベクトル束**という．

(i) $I^2 = id$ かつ $I \neq id$;

(ii) (I) の条件より I は固有値 ± 1 を持つ．固有値 $+1$ の固有空間を E^+，固有値 -1 の固有空間を E^- とする．このとき E^+ と E^- はどちらも n 階のベクトル束であり，$E = E^+ \oplus E^-$ が成り立つ；

(iii) $\tau(I(\eta), \zeta) + \tau(\eta, I(\zeta)) = 0$. ただし，$\eta, \zeta \in \Gamma(E)$.

注意 3.2 ω を $\omega(\eta, \zeta) := \tau(\eta, I(\zeta))$ で定めると，これは E 上のシンプレクティック形式になっていることが条件 (iii) からわかる．さらに，$\tau|_{E^+} = 0$ かつ $\tau|_{E^-} = 0$ であることもわかる．

例 3.3 $E^+ := TM$，$E^- := T^*M$，$E := E^+ \oplus E^-$ とする．ただし T^*M は余接ベクトル束である．E 上の計量 τ と束写像 $I : E \to E$ を

$$\tau(\eta^+ \oplus \eta^-, \zeta^+ \oplus \zeta^-) := \zeta^-(\eta^+) + \eta^-(\zeta^+),$$
$$I(\eta^+ \oplus \eta^-) := \eta^+ \oplus -\eta^-$$

で定義する．ただし，$\eta^+, \zeta^+ \in \Gamma(E^+)$，$\eta^-, \zeta^- \in \Gamma(E^-)$. このとき，$(E, \tau, I)$ はパラエルミートベクトル束になっていることが確かめられる．

定義 3.4 ∇ を E^+ 上の接続とする．双対接続 ∇^* を E^- 上の接続として次の方程式を満たすように定義する．

$$X\tau(\eta^+, \eta^-) = \tau(\nabla_X \eta^+, \eta^-) + \tau(\eta^+, \nabla^*_X \eta^-).$$

ただし，$\eta^+ \in \Gamma(E^+)$，$\eta^- \in \Gamma(E^-)$.

注意 3.2 と τ の非退化性より ∇^* は一意に存在し，明らかに $(\nabla^*)^* = \nabla$ が成り立つ．以下ではわかりやすく，∇ を ∇^+，∇^* を ∇^- と表現する．(E, τ, I) をパラエルミートベクトル束とし，束写像 $\Phi : \Gamma(TM) \to \Gamma(E)$ を各点 $p \in M$ で $\mathrm{Im}\Phi_p \subset E_p$ が ω に関して Lagranage 部分空間になっている (i.e., $rank\Phi_p = n$ かつ $\mathrm{Im}\Phi_p$ 上で ω が消えている) ものとしてとる．また，$\pi^+ : E \to E^+$ と $\pi^- : E \to E^-$ をそれぞれ射影とし，$\Phi^+ := \pi^+ \circ \Phi$ と $\Phi^- := \pi^- \circ \Phi$ を定義する．さらに，

$$\xi^+ := \Phi^+(X), \quad \eta^+ := \Phi^+(Y), \quad \zeta^+ := \Phi^+(Z),$$
$$\xi^- := \Phi^-(X), \quad \eta^- := \Phi^-(Y), \quad \zeta^- := \Phi^-(Z).$$

と記号を置く．

$$\begin{tikzcd}
(E^+, \nabla^+) \arrow[dr] & TM \arrow[l, "\Phi^+"'] \arrow[r, "\Phi^-"] \arrow[d] & (E^-, \nabla^-) \arrow[dl] \\
& M &
\end{tikzcd}$$

定義 3.5 $(0, 2)$ テンソル h を次で定義する．

$$h(X, Y) := \tau(\Phi(X), \Phi(Y)) = 2\tau(\xi^+, \eta^-).$$

最後の等式は注意 3.2 と Φ の仮定から成り立つ．

注意 3.6 h が非退化であることの必要十分条件は Φ^+ と Φ^- がどちらも同型であることである．またこのとき，TM 上の接続として $\widetilde{\nabla}^+$ と $\widetilde{\nabla}^-$ を

$$\Phi^+(\widetilde{\nabla}^+_X Y) := \nabla^+_X \eta^+, \qquad \Phi^-(\widetilde{\nabla}^-_X Y) := \nabla^-_X \eta^- \tag{3.1}$$

で定めることができる．∇^+ と ∇^- の双対性より，$\widetilde{\nabla}^+$ と $\widetilde{\nabla}^-$ も h に関して双対性を持つことがわかる．つまり，

$$Xh(Y,Z) = h(\widetilde{\nabla}^+_X Y, Z) + h(Y, \widetilde{\nabla}^-_X Z)$$

が満たされる．

定義 3.7 (cf. [9]) ∇^+ が *relatively torsion-free* とは

$$\nabla^+_X \eta^+ - \nabla^+_Y \xi^+ - \Phi^+([X,Y]) = 0$$

の条件を満たすときのことを言う．同様に ∇^- が relatively torsion-free とは

$$\nabla^-_X \eta^- - \nabla^-_Y \xi^- - \Phi^-([X,Y]) = 0$$

を満たすときを言う．

注意 3.8 h が非退化であるとき，∇^+ (resp. ∇^-) が relatively torsion-free であるとき，(3.1) によって定義される接続 $\widetilde{\nabla}^+$ (resp. $\widetilde{\nabla}^-$) は torsion-free になっていることが確かめられる．

定義 3.9 3 次テンソル C を次で定義する．

$$C(X,Y,Z) := -2\{\tau(\nabla^+_X \eta^+, \zeta^-) - \tau(\zeta^+, \nabla^-_X \eta^-)\}.$$

注意 3.10 h が非退化であるとき，定義 3.9 の 3 次テンソル C について，$C = \widetilde{\nabla}^+ h$ が成り立つ．

定義 3.11 [3, 4] *Kossowski* **擬接続** $\Gamma : \Gamma(TM) \times \Gamma(TM) \times \Gamma(TM) \to C^\infty(M)$ を次で定義する．

$$\begin{aligned}\Gamma(X,Y,Z) = X\tau(\eta^+, \zeta^-) + Y\tau(\zeta^+, \xi^-) - Z\tau(\xi^+, \eta^-) \\ - \tau(\xi^+, \Phi^-([Y,Z])) + \tau(\eta^+, \Phi^-([Z,X])) + \tau(\zeta^+, \Phi^-([X,Y])).\end{aligned}$$

注意 3.12 h が非退化であるとき，Kossowski 擬接続 Γ は Levi-Civita 接続 ∇^h を用いて

$$h(\nabla^h_X Y, Z) = \Gamma(X,Y,Z)$$

と表現できる．

計量が退化している場合における命題 2.2 の拡張として次の 4 つの条件を考える．

(i) ∇^+ は relatively torsion-free;

(ii) ∇^- は relatively torsion-free;

(iii) C は全対称 (i.e., $C(X,Y,Z) = C(Y,X,Z) = C(X,Z,Y)$);

(iv) $\Gamma(X,Y,Z) = \tau(\nabla^+_X \eta^+, \zeta^-) + \tau(\zeta^+, \nabla^-_X \eta^-)$.

命題 3.13 上記の 4 つの条件に関して次が成り立つ.

(1) (i) と (ii) が成り立つとき, (iii) と (iv) も同時に成り立つ;

(2) Φ^+ が同型かつ (i) が成り立つとする. このとき (iii) または (iv) が成り立つとき, (ii) も成り立つ;

(3) Φ^- が同型かつ (ii) が成り立つとする. このとき (iii) または (iv) が成り立つとき, (i) も成り立つ;

(4) h が非退化であるとする. このとき, いずれかの 2 つの条件を満たせば残りの 2 つの条件も同時に満たされる.

統計構造は 4 つの条件の内のいずれか 2 つを満たすものとして定義したが, 計量が退化している場合には次で統計構造の拡張を定義する.

定義 3.14 ∇^+ と ∇^- がどちらも relatively torsion-free であるとき, 組 $(h, (E, \tau, I), \Phi, \nabla^+, \nabla^-)$ を M 上の**概コダッチ構造**と呼ぶ.

注意 3.15 h が非退化であるとき, 上記の 4 つの条件は命題 2.2 の 4 つの条件と同じことを表している. したがって, h が非退化な概コダッチ構造 $(h, (E, \tau, I), \Phi, \nabla^+, \nabla^-)$ が与えられたとき, (3.1) によって定義される $\widetilde{\nabla}^+$ と $\widetilde{\nabla}^-$ を考えることで通常の意味の統計構造 $(h, \widetilde{\nabla}^+, \widetilde{\nabla}^-)$ を誘導することができる.

例 3.16 $(h, \widetilde{\nabla}^+, \widetilde{\nabla}^-)$ を M 上の統計構造とし, (E, τ, I) を例 3.3 で構成された M 上のパラエルミートベクトル束とする. 束写像 $\Phi = \Phi^+ \oplus \Phi^- : TM \to E$ を

$$\Phi^+(X) := id(X), \qquad \Phi^-(X) := h(X, -)$$

で定め, E^+ と E^- 上の接続 ∇^+ と ∇^- を (3.1) によって定義する. このとき, $(h, (E, \tau, I), \Phi, \nabla^+ \nabla^-)$ は概コダッチ構造になっている.

上述によって計量が退化する場合における統計構造を定義した. Nakajima-Ohmoto([8]) の先行研究によって, 計量が退化する場合における双対平坦構造が**概ヘッセ構造**として拡張されている. 詳細は解説しないが, 概ヘッセ構造と概コダッチ構造の定義は少し異なる. しかし次の事実が成り立つ.

定理 3.17 概ヘッセ構造が与えられたとき, ∇^+ と ∇^- が共に平坦な概コダッチ構造が自然に誘導することができる. また逆に, ∇^+ と ∇^- が共に平坦な概コダッチ構造が与えられたとき, 概ヘッセ構造が自然に誘導される.

参考文献

[1] S. Amari, *Information Geometry and Its Application*, Applied Math. Sci., 194, Springer (2016).

[2] S. Amari, H. Nagaoka, *Method of Information Geometry*, A.M.S., Oxford Univ. Press (2000).

[3] M. Hasegawa, A. Honda, K. Naokawa, K. Saji, M. Umehara, K. Yamada, Intrinsic properties of surface with singularities, International J. Math., Vol.26, No.4, 1540008, (2015).

[4] M. Kossowski, Realizing a singular first fundamental form as a nonimmersed surface in Euclidian 3-space, J. Geom. **81** (2004), 101–113.

[5] K. Kayo, Statistical manifold with degenerate metric, preprint, arXive:2310.18599.

[6] M.-A. Lawn, L. Schäfer, Decompositions of para-complex vector bundles and para-complex affine immersions, Result. Math. **48** (2005), 246–274.

[7] H. Matsuzoe, Statistical manifolds and affine differential geometry, Advanced Studies in Pure Mathematics **57** (2010), 303–321.

[8] N. Nakajima, T. Ohmoto, The dually flat structure for singular models, Information Geometry. J. **4** (2021), 31–64.

[9] K. Saji, M. Umehara, K. Yamada, *Differential Geometry of Curves and Surface with Singularities* (in Japanese), Maruzen Shuppan, Co. Ltd. (2017).