情報幾何学の基礎

Chapter 5.2 Fisher 計量と α 接続
Chapter 5.3 指数型接続と混合型接続

田淵智久

2022/1/6

早稲田大学基幹理工学部情報通信学科

目次

前回までの流れと今回の目的

前回までの流れと今回の目的

- 5 章では, 有限集合上の確率分布全体からなる空間がどのような幾何構造になっているかを見ていく.
- 前回の 5.1 節では, 確率分布の幾何構造は, 「全ての事象における確率の和が 1 となる」という確率の性質から, 確率シンプレックスという凸集合になっていることを確認した.
- そして, このような確率分布の幾何構造に要請されることとして, ラベルの取り替えには不変であって欲しいっていうがあった. このことと, Sn-1 の幾何構造は, それよりも次元の大きい Sl-1 の幾何構造から誘導されるだろう, 誘導されるべきだという考え方から, マルコフ埋め込みを導入しました. このマルコフ埋め込みという写像において, 不変となるような計量がフィッシャー計量になることを示した.
- 今回は確率分布の幾何構造をもっと詳しく見ていく. 具体的には, フィッシャー計量 g に付随する接続がどんなものであるかを見ていく. そしてそれが α-接続であることを示して, その後は, 応用上よく登場する α が +-1 の場合について調べる.nabla+1 と nabla-1 は双対だねとか, (g, nabla+1, nabla-1) ば双対平坦だねとか見ていく. 最後に, この接続のもとで, 平行移動がどう定義されてるのかを調べる.

Fisher 計量と α 接続

Markov 埋め込み f に関する $g^{[n]}$ の不変性

$\mathcal{S}_{n-1}$ 上の計量 $g^{[n]}$ が Markov 埋め込み $f : \mathcal{S}_{n-1} \mapsto \mathcal{S}_{l-1}$ に関して不変であるとは

$$g_p^{[n]}(X, Y) = g_{f(p)}^{[l]}(f_* X, f_* Y) \tag{5.6}$$

が言えることである. これについては 5.1 節で示した.

定理 5.1.1 Chentsov (復習)

$\mathcal{S}_{n-1}$ 上の (0,2) 型テンソル場 $g^{[n]}$ からなる列 $\{g^{[n]}; n = 2, 3, \ldots\}$ であって, 任意の Markov 埋め込み f に関する不変性

$$g_p^{[n]}(X, Y) = g_{f(p)}^{[l]}(f_* X, f_* Y)$$

を満たすものは, 定数倍を除いて

$$g_p^{[n]}(X, Y) = \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega))$$

に限られる.

Markov 埋め込み f に関する $\nabla^{[n]}$ の不変性

$\mathcal{S}_{n-1}$ 上のアファイン接続 $\nabla^{[n]}$ が Markov 埋め込み $f : \mathcal{S}_{n-1} \mapsto \mathcal{S}_{l-1}$ に関して不変であるとは

$$g_p^{[n]}(\nabla_X^{[n]} Y, Z) = g_{f(p)}^{[l]}(\nabla_{f_* X}^{[l]} Y, f_* Z) \tag{5.7}$$

が言えることである. 前ページで確認した Markov 埋め込み f に不変な計量 $g^{[n]}$ に付随する Riemann 接続を $\overline{\nabla}^{[n]}$ と書くと, 式 (5.6) から $\overline{\nabla}^{[n]}$ が式 (5.7) を満たすことが示せる.

定義 (復習 p.79)

$$g(\nabla_{\partial_i}\partial_j, \partial_k) = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}) \tag{3.25}$$

f に関する $g^{[n]}$ の不変性に注意すると, 以下のように $\overline{\nabla}^{[n]}$ の不変性が求まる.

$$\begin{aligned}
g_{f(p)}^{[l]}(\overline{\nabla}_{f_*\partial_i}^{[n]} f_*\partial_j, f_*\partial_k) &= \frac{1}{2}\left\{(f_*\partial_i) g_{f(p)}^{[l]}(f_*\partial_j, f_*\partial_k) + (f_*\partial_j) g_{f(p)}^{[l]}(f_*\partial_k, f_*\partial_i) - (f_*\partial_k) g_{f(p)}^{[l]}(f_*\partial_i, f_*\partial_j)\right\} \\
&= \frac{1}{2}\left\{\partial_i g_p^{[n]}(\partial_j, \partial_k) + \partial_j g_p^{[n]}(\partial_k, \partial_i) - \partial_k g_p^{[n]}(\partial_i, \partial_j)\right\} \\
&= g_p^{[n]}(\overline{\nabla}_{\partial\ i}^{[n]} \partial_j, \partial_k)
\end{aligned}$$

従って, Riemann 接続 $\overline{\nabla}^{[n]}$ は $\mathcal{S}_{n-1}$ 上に許容される接続の一つである.
ここから, $\mathcal{S}_{n-1}$ 上に許容される接続は他にないか考える. 多様体上に共変微分 $\overline{\nabla}^{[n]}$ を 1 つ固定する. 3.1 節より,

$$\overline{\nabla}^{[n]}_X Y$$

はテンソル性を満たさない. しかし, 共変微分を任意に 2 つ, $\overline{\nabla}^{[n]}$ と $\nabla^{[n]}$ を選ぶと,

$$\nabla^{[n]}_X Y - \overline{\nabla}^{[n]}_X Y$$

はテンソル性を満たし, (1,2) 型テンソル場となる. $(\mathcal{X} \times \mathcal{X} \to \mathcal{X})$
そしてこれを計量 $g^{[n]}_p$ で射影すると,

$$g^{[n]}_p(\nabla^{[n]}_X Y, Z) - g^{[n]}_p(\overline{\nabla}^{[n]}_X Y, Z)$$

は (0,3) 型テンソル場となる. $(\mathcal{X} \times \mathcal{X} \times \mathcal{X} \to C^\infty)$

以上から, $\nabla^{[n]}$ が Markov 埋め込み f に関して不変であるとは, 上の (0,3) 型テンソル場が Markov 埋め込み f に関して不変であることに他ならない. 定理 5.1.2 より, Markov 埋め込み f に関する不変性を満たす (0,3) 型テンソル場は定数倍を除いて

$$S^{[n]}_p(X, Y, Z) = \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega))(Z \log p(\omega)) \tag{5.5}$$

と定まることが言える.

定理 5.1.2 Chentsov (復習)

$\mathcal{S}_{n-1}$ 上の (0,3) 型テンソル場 $S^{[n]}$ からなる列 $\{S^{[n]}; n = 2, 3, \ldots\}$ であって, 任意の Markov 埋め込み f に関する不変性

$$S^{[n]}_p(X, Y, Z) = S^{[l]}_{f(p)}(f_* X, f_* Y, f_* Z) \tag{5.4}$$

を満たすものは, 定数倍を除いて

$$S^{[n]}_p(X, Y, Z) = \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega))(Z \log p(\omega)) \tag{5.5}$$

に限られる.

定数倍の部分を $-\alpha/2$ とおくと,

$$g^{[n]}_p(\nabla^{[n]}_X Y, Z) - g^{[n]}_p(\overline{\nabla}^{[n]}_X Y, Z) = -\frac{\alpha}{2} S^{[n]}_p(X, Y, Z)$$

と表せる. 以上から, Markov 埋め込み f に関する不変性 (5.7) を満たす接続 $\nabla^{[n]}$ がこの式によって 1 つに定まることがわかる. 以上をまとめて, 次の結論を得る.

Chentsov の定理

定理 5.2.1. (Chentsov)

Markov 埋め込みのもとでの不変性 (5.6) を満たす S_{n-1} 上の Riemann 計量は, 定数倍を除いて

$$g_p(X,Y) := \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega)) \tag{5.8}$$

に限られる. 一方, 不変性 (5.7) を満たすファイン接続 $\nabla^{(\alpha)}$ は, 関係

$$g_p(\nabla^{(\alpha)}_X Y, Z) := g_p(\overline{\nabla}_X Y, Z) - \frac{\alpha}{2} S_p(X,Y,Z) \tag{5.9}$$

により, 実数 α と 1 対 1 に対応する. ただし, $\overline{\nabla}$ は計量 (5.8) に付随する Riemann 接続, S_p は

$$S_p(X,Y,Z) := \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega))(Z \log p(\omega))$$

で定義される (0,3) 型対称テンソル場である.

Fisher 計量と α 接続

定義

$$g_p(X, Y) := \sum_{\omega=1}^{n} p(\omega)(X \log p(\omega))(Y \log p(\omega)) \tag{5.8}$$

$$g_p(\nabla^{(\alpha)}_X Y, Z) := g_p(\overline{\nabla}_X Y, Z) - \frac{\alpha}{2} S_p(X, Y, Z) \tag{5.9}$$

(5.8) で定まる計量 g を S_{n-1} のFisher 計量といい, 各実数 α に対し (5.9) で定まる接続 $\nabla^{(\alpha)}$ を S_{n-1} のα-接続という.

特に, $\nabla^{(0)}$ は計量 g に付随した Riemann 接続である.

定理 3.6.2 (復習)

Riemann 多様体上 (M, g) のアファイン接続 ∇ が Riemann 接続であるための必要十分条件は, ∇ が計量的であって, かつ捩率がゼロとなることである.

指数型接続と混合型接続

$\mathcal{S}_{n-1}$ はどのような多様体か?

有限確率分布空間 $\mathcal{S}_{n-1}$ が有する極めて自然な不変性の要請から, $\mathcal{S}_{n-1}$ 上に許容される計量と接続がそれぞれ Fisher 計量と α-接続に限られることが Chentsov の定理で明らかになった. では, これらの幾何構造を有する空間 $\mathcal{S}_{n-1}$ はどのような多様体になっているかを調べていく.

定理 5.3.1.

有限確率分布空間 S_{n-1} の Fisher 計量を g, α 接続を $\nabla^{(\alpha)}$ とする.

(i) 各 $\alpha \in \mathbb{R}$ に対し, $\nabla^{(\alpha)}$ の捩率はゼロである.

(ii) 各 $\alpha \in \mathbb{R}$ に対し, $\nabla^{(\alpha)}$ と $\nabla^{(-\alpha)}$ は g に関して互いに双対

(iii) $\mathcal{S}_{n-1}$ は双対構造 $(g, \nabla^{(-1)}, \nabla^{(+1)})$ に関して双対平坦

証明

$\nabla^{(\alpha)}$ の捩率が 0 であることを示す際に Riemann 計量 g が持つ次の性質を用いる.

「全ての $Z \in \mathcal{X}(M)$ に対して, $g(Y, Z) = 0$ ならば, $Y = 0$ である.」

よって, $g(T(X, Y), Z) = 0$ であることを示すことによって $T = 0$ であることを示す.
$\nabla^{(0)}$ は Riemann 接続であって捩率が 0 であることを用いれば,

$$
\begin{aligned}
g(T(X, Y), Z) &= g(\nabla^{(\alpha)}_X Y - \nabla^{(\alpha)}_Y X - [X, Y], Z) \\
&= g(\nabla^{(0)}_X Y - \nabla^{(0)}_Y X - [X, Y], Z) - \frac{\alpha}{2}\{S(X, Y, Z) - S(Y, X, Z)\} \\
&= 0
\end{aligned}
$$

となって, (i) が証明される.

定義 (復習)

アファイン接続 ∇ を持つ多様体 M 上で定義された 3 重 $C^\infty(M)$-線形写像

$$T \quad : \quad \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M) : (X, Y) \longmapsto \nabla_X Y - \nabla_Y X - [X, Y]$$

を接続 ∇ の捩率テンソル場という.

定義 (復習)

アフィン接続 ∇ を持つ Riemann 多様体 (M,g) において

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z) \tag{4.1}$$

で定義されるアフィン接続 ∇^* を計量 g に関する ∇ の双対アフィン接続という.

また, 定理 5.2.1. より,

$$g_p(\nabla^{(\alpha)}_X Y, Z) := g_p(\nabla^{(0)}_X Y, Z) - \frac{\alpha}{2} S_p(X, Y, Z)$$

だから, $g_p(\nabla^{(0)}_X Y, Z)$ は次のように書き直せる.

$$\begin{aligned} g_p(\nabla^{(0)}_X Y, Z) &= g_p(\nabla^{(\alpha)}_X Y, Z) + \frac{\alpha}{2} S_p(X, Y, Z) \\ g_p(\nabla^{(0)}_X Y, Z) &= g_p(\nabla^{(-\alpha)}_X Y, Z) - \frac{\alpha}{2} S_p(X, Y, Z) \end{aligned}$$

これと, S が (0,3) 型対称テンソル場であることに注意すると, (ii) が証明される.

$$\begin{aligned} Xg(Y,Z) &= g(\nabla^{(0)}_X Y, Z) + g(Y, \nabla^{(0)}_X Z) \\ &= g(\nabla^{(\alpha)}_X Y, Z) + g(Y, \nabla^{(-\alpha)}_X Z) + \frac{\alpha}{2}\{S(X,Y,Z) - S(X,Z,Y)\} \\ &= g(\nabla^{(\alpha)}_X Y, Z) + g(Y, \nabla^{(-\alpha)}_X Z) \end{aligned}$$

(iii) を示すには, S_{n-1} が $\nabla^{(-1)}$-アファイン座標系を持つことを言えばよい.

定義 (復習)

双対構造 (g, ∇, ∇^*) を持つ多様体 M において, ∇ に関する曲率 R も捩率 T も共にゼロとなり, かつ ∇^* に関する曲率 R^* も捩率 T^* も共にゼロとなるとき, M は双対平坦であるという.

定理 4.1.2 (復習)

∇-曲率 R がゼロであることと ∇-曲率 R^* がゼロであることは同値である.

定義 (復習)

アファイン接続 ∇ を持つ多様体 M の座標近傍 $(U; x^1, \ldots, x^n)$ において, 接続係数 $\{\Gamma_{ij}{}^k\}$ が全て恒等的にゼロとなるとき, $(U; x^1, \ldots, x^n)$ を∇-アファイン座標近傍といい, 局所座標系 (x^i) を∇-アファイン座標系という.

定理 3.5.1 (復習)

アファイン接続 ∇ を持つ多様体 M において, 次の 2 条件は同値である.

(i) M は ∇-平坦である.

(ii) M の各点の周りに ∇-アファイン座標近傍が存在する.

(iii) を示すには $\mathcal{S}_{n-1}$ が $\nabla^{(-1)}$-アファイン座標系を持つことを言えばよい?

$\mathcal{S}_{n-1}$ が $\nabla^{(-1)}$-アファイン座標系を持つ

$\longrightarrow$ $\nabla^{(-1)}$-曲率 R と $\nabla^{(-1)}$-捩率 T がゼロ

$\longrightarrow$ $\nabla^{(-1)}$-曲率 R と $\nabla^{(+1)}$-曲率 R^* と $\nabla^{(-1)}$-捩率 T がゼロ

$\nabla^{(\alpha)}$ の捩率がゼロであることは先程示したので,

$\longrightarrow$ $\nabla^{(-1)}$-曲率 R と $\nabla^{(+1)}$-曲率 R^* と $\nabla^{(-1)}$-捩率 T と $\nabla^{(+1)}$-捩率 T^* がゼロ

$\longrightarrow$ $\mathcal{S}_{n-1}$ は双対構造 $(g, \nabla^{(-1)}, \nabla^{(+1)})$ に関して双対平坦

従って, $\mathcal{S}_{n-1}$ が $\nabla^{(-1)}$-アファイン座標系を持つことが言えれば, (iii) が示せる.

$\mathcal{S}_{n-1}$ が $\nabla^{(-1)}$-アファイン座標系を持つことを示す

$\mathcal{S}_{n-1}$ の一般の座標系 (x^i) に関する α 接続 $\nabla^{(\alpha)}$ の接続係数は

$$\begin{aligned}\Gamma^{(\alpha)}_{ij,k} &\coloneqq g(\nabla^{(\alpha)}_{\partial_i}\partial_j, \partial_k)\\ &= \frac{1-\alpha}{2}S_{ijk} + \sum_{\omega=1}^{n} p(\omega)(\partial_i\partial_j \log p(\omega))(\partial_k \log p(\omega))\end{aligned}$$

となる. ただし, $S_{ijk} \coloneqq S(\partial_i, \partial_j, \partial_k)$ である. まず, これについて示す.

$$\begin{aligned}\Gamma^{(\alpha)}_{ij,k} &= g(\nabla^{(\alpha)}_{\partial_i}\partial_j, \partial_k)\\ &= g(\nabla^{(0)}_{\partial_i}\partial_j, \partial_k) - \frac{\alpha}{2}S_{ijk}\\ &= \Gamma^{(0)}_{ij,k} - \frac{\alpha}{2}S_{ijk}\end{aligned}$$

となる. $\Gamma^{(0)}_{ij,k}$ について展開する. 以降の計算は補足資料に掲載.

$\alpha = -1$ のとき,

$$
\begin{aligned}
\Gamma^{(-1)}_{ij,k} &= S_{ijk} + \sum_{\omega=1}^{n} p(\omega)(\partial_i \partial_j \log p(\omega))(\partial_k \log p(\omega)) \\
&= \sum_{\omega=1}^{n} p(\omega)(\partial_i \log p(\omega))(\partial_j \log p(\omega))(\partial_k \log p(\omega)) \\
&\quad + \sum_{\omega=1}^{n} p(\omega)(\partial_i \partial_j \log p(\omega))(\partial_k \log p(\omega)) \\
&= \sum_{\omega=1}^{n} p(\omega) \left\{ \left(\frac{\partial_i p(\omega)}{p(\omega)} \right) \left(\frac{\partial_j p(\omega)}{p(\omega)} \right) + \partial_i \left(\frac{\partial_j p(\omega)}{p(\omega)} \right) \right\} (\partial_k \log p(\omega)) \\
&= \sum_{\omega=1}^{n} (\partial_i \partial_j p(\omega))(\partial_k \log p(\omega))
\end{aligned}
$$

式 (5,3) で導入した $\mathcal{S}_{n-1}$ の座標系 (ξ^a), すなわち,

$$
p(\omega) = \sum_{a=1}^{n-1} \xi^a \delta_a(\omega) + \left(1 - \sum_{a=1}^{n-1} \xi^a \right) \delta_n(\omega)
$$

に対しては $\partial_a \partial_b p(\omega) = 0$ となるので, $\Gamma^{(-1)}_{ab,c} = 0$ が示された. □

指数型接続と混合型接続

定義

$\alpha = -1$ に対応する接続 $\nabla^{(-1)}$ を混合型接続といい, 記号 $\nabla^{(m)}$ で表す. 一方, $\alpha = +1$ に対応する接続 ∇^{+1} を指数型接続といい, 記号 $\nabla^{(e)}$ で表す.

双対構造 $(g, \nabla^{(m)}, \nabla^{(e)})$ に関して双対平坦な多様体 $\mathcal{S}_{n-1}$ の性質を調べていく. これまでの議論において, 接ベクトル $X \in T_pS$ はしばしば p や $\log p$ に作用する形で登場した. そこでこれらを, 接ベクトル X の確率変数による表現とみなし, 以下の定義を行う.

双対平坦な多様体 S_{n-1} の性質

定義

各 $X \in T_pS$ に対し,

$$Xp := (Xp(1), \ldots, Xp(n))$$

で定まる確率変数 Xp を X の混合型表現もしくは簡単にm-表現という. そして, その全体を接空間 T_pS の m-表現といい, $T_p^{(m)}S$ で表す. すなわち,

$$T_p^{(m)}S =: \{Xp; X \in T_pS\}$$

一方, 各 $X \in T_pS$ に対し,

$$Xp \log p := (X \log p(1), \ldots, X \log p(n))$$

で定まる確率変数 $X \log p$ を X の指数型表現もしくは簡単にe-表現という. そして, その全体を接空間 T_pS の e-表現といい, $T_p^{(e)}S$ で表す. すなわち,

$$T_p^{(e)}S := \{X \log p; X \in T_pS\}$$

接空間 T_pS の m-表現と e-表現

定理 5.3.2

接空間 T_pS の m-表現, e-表現は, それぞれ次のように特徴づけられる.

$$T_p^{(m)}S = \left\{ f : \Omega \to \mathbb{R}; \sum_{\omega\in\Omega} f(\omega) = 0 \right\}$$

$$T_p^{(e)}S = \left\{ f : \Omega \to \mathbb{R}; \sum_{\omega\in\Omega} p(\omega)f(\omega) = 0 \right\}$$

(証明)　上式を示す. 左辺 $\subset$ 右辺は

$$\sum_{\omega\in\Omega} Xp(\omega) = X\left(\sum_{\omega\in\Omega} p(\omega)\right) = X1 = 0$$

から分かる. そして両辺共に n-1 次元線形空間であることから, 左辺 = 右辺が言える.

次に下式を示す. 左辺 ⊂ 右辺は

$$\sum_{\omega\in\Omega} p(\omega)(X\log p(\omega)) = \sum_{\omega\in\Omega} p(\omega)\left(\frac{Xp(\omega)}{p(\omega)}\right) = X\left(\sum_{\omega\in\Omega} p(\omega)\right) = X1 = 0$$

から分かる. そして両辺共に n-1 次元線形空間であることから, 左辺 = 右辺がわかる. □

双対性の確認

式 (5.13) に登場する

$$\sum_{\omega\in\Omega} p(\omega)f(\omega)$$

は確率分布 p のもとでの確率変数 f の期待値である. これを $E_p[f]$ と書くことにする. Fisher 計量 g は次のように書ける.

$$\begin{aligned} g(X,Y) &= \sum_{\omega\in\Omega} p(\omega)(X\log p(\omega))(Y\log p(\omega)) \\ &= \sum_{\omega\in\Omega} p(\omega)\left(\frac{Xp(\omega)}{p(\omega)}\right)(Y\log p(\omega)) \\ &= \sum_{\omega\in\Omega}(Xp(\omega))(Y\log p(\omega)) \end{aligned}$$

また,

$$g_p(\nabla^{(\alpha)}_X Y, Z) = \frac{1-\alpha}{2} S(X,Y,Z) + \sum_{\omega\in\Omega}(XY\log p(\omega))(Z\log p(\omega))$$

であるから,

$$g_p(\nabla^{(e)}_X Y, Z) = \sum_{\omega\in\Omega}(XY \log p(\omega))(Z \log p(\omega)) \tag{5.15}$$

$$g_p(\nabla^{(m)}_X Y, Z) = \sum_{\omega\in\Omega}(XY p(\omega))(Z \log p(\omega)) \tag{5.16}$$

と書ける. これらを用いると,

$$\begin{aligned} Xg(Y,Z) &= X\left\{\sum_{\omega\in\Omega}(Y p(\omega))(Z \log p(\omega))\right\} \\ &= \sum_{\omega\in\Omega}(XY p(\omega))(Z \log p(\omega)) + \sum_{\omega\in\Omega}(Y p(\omega))(XZ \log p(\omega)) \\ &= g_p(\nabla^{(m)}_X Y, Z) + g_p(Y, \nabla^{(e)}_X Z) \end{aligned}$$

このように, 接続の双対性が機械的に計算することで確認できる.

m-平行移動と e-平行移動の定義

S は双対構造 $(g, \nabla^{(m)}, \nabla^{(e)})$ に関して双対平坦であった. $\nabla^{(m)}$-曲率も $\nabla^{(e)}$-曲率もゼロであるから, 3.3 節より, どちらの接続に関する平行移動も, 始点と終点によってのみ決まり, 経路によらない. $\nabla^{(m)}$ に関する平行移動, $\nabla^{(e)}$ に関する平行移動がどのようなものかについて見ていく.

定義

2 点 $p, q \in S$ に対し,

$$\prod^{(m)} : T_p^{(m)} S \longrightarrow T_q^{(m)} S : f \longrightarrow f$$

で定まる遠隔平行移動を混合型平行移動もしくは簡単にm-平行移動という. また,

$$\prod^{(e)} : T_p^{(e)} S \longrightarrow T_q^{(e)} S : f \longrightarrow f - E_q[f]$$

で定まる遠隔平行移動を指数型平行移動もしくは簡単にe-平行移動という.

式 (5.12) より, 接空間 T_pS の m-表現は p に依存しない. よって, 確率変数 f を p から q へ移動する際, 何もせずにそのまま移動して, T_pS の元の m-表現と見なすことができる. 一方, 式 (5.13) より, 接空間 T_pS の e-表現は p に依存する. p での期待値がゼロになる確率変数だけが p での接ベクトルの e-表現と見なすことができる.

接続と平行移動の関係

m-平行移動と e-平行移動の無限小版がそれぞれ $\nabla^{(m)}$ と $\nabla^{(e)}$ に一致する.

定理 5.3.3

$t=0$ で $p_0 \in S$ を通る $X \in \chi(S)$ の積分曲線を p_t とする. すなわち $\dot{p}_t = X_p$ である. 任意のベクトル場 $Y \in \chi(S)$ に対し,

$$\left(\nabla_X^{(m)} Y\right)_{p_0} = \lim_{t \to 0} \frac{\prod^{(m)}{}_0^t (Y_{p_t}) - Y_{p_0}}{t} \tag{5.17}$$

および

$$\left(\nabla_X^{(e)} Y\right)_{p_0} = \lim_{t \to 0} \frac{\prod^{(e)}{}_0^t (Y_{p_t}) - Y_{p_0}}{t} \tag{5.18}$$

が成り立つ.

証明

式 (5.17) を示す. 右辺で極限をとる前の式の m-表現は以下の確率変数である.

$$\frac{\prod^{(m)t}{}_0(Y_{p_t}) - Y_{p_0}}{t}$$

任意の $Z \in \mathcal{X}(\mathcal{S})$ に対し,

$$\begin{aligned}
& \lim_{t\to 0} \sum_{\omega\in\Omega} \left(\frac{\prod^{(m)t}{}_0(Y_{p_t} p_t(\omega)) - Y_{p_0} p_0(\omega)}{t} \right) (Z_{p_0} \log p_0(\omega)) \\
= \quad & \lim_{t\to 0} \sum_{\omega\in\Omega} \left(\frac{(Y_{p_t} p_t(\omega)) - Y_{p_0} p_0(\omega)}{t} \right) (Z_{p_0} \log p_0(\omega)) \\
= \quad & \sum_{\omega\in\Omega} p(\omega)(XY p(\omega))(Z \log p(\omega)) \Big|_{p=p_0} \\
= \quad & g_p(\nabla^{(m)}_X Y, Z)
\end{aligned}$$

定義 (復習)

$X \in \mathcal{X}(M)$ に対し, C^∞ 曲線 $C = \{p(t); a < t < b\}$ であって, その各点 $p(t)$ における速度ベクトル $\dot{p}(t)$ が $X_{p(t)}$ に一致するようなものを X の積分曲線という.

証明

式 (5.18) を示す. 右辺で極限をとる前の式の e-表現は以下の確率変数である.

$$\frac{\prod^{(e)t}{}_0(Y_{p_t}) - Y_{p_0}}{t}$$

任意の $Z \in \mathcal{X}(\mathcal{S})$ に対し,

$$\begin{aligned}
& \lim_{t\to 0}\sum_{\omega\in\Omega}\left(\frac{\prod^{(e)t}{}_0(Y_{p_t}\log p_t(\omega)) - Y_{p_0}\log p_0(\omega)}{t}\right)(Z_{p_0}\log p_0(\omega)) \\
= \quad & \lim_{t\to 0}\sum_{\omega\in\Omega}\left(\frac{(Y_{p_t}\log p_t(\omega) - E_{p_0}[Y_{p_t}\log p_t]) - Y_{p_0}\log p_0(\omega)}{t}\right)(Z_{p_0}\log p_0(\omega)) \\
= \quad & \lim_{t\to 0}\sum_{\omega\in\Omega}\left(\frac{Y_{p_t}\log p_t(\omega) - Y_{p_0}\log p_0(\omega)}{t}\right)(Z_{p_0}\log p_0(\omega)) \\
= \quad & \sum_{\omega\in\Omega} p(\omega)(XY\log\log p(\omega))(Z\log p(\omega))\Big|_{p=p_0} \\
= \quad & g_p(\nabla^{(e)}_X Y, Z)
\end{aligned}$$

□

参考文献

[1] 藤原彰夫, 情報幾何学の基礎, 共立出版, 2021