# 2019数学アドベントカレンダー

2019年12月16日　　15:27

今回は最近読んでいる情報幾何の本で学んだことを説明しようと思います。
~~Texで書ききる体力がなかったため、~~実験的にOneNoteにまとめました。

情報幾何は確率分布を幾何的な対象だと捉えます。
例えば、じゃんけんをする時、グーを確率p、パーを確率q、
チョキを確率rで出すとします。
これを3次元上の点(p,q,r)に対応させます。
確率はp+q+r=1,p,q,r全て0以上という条件から、
この条件を満たす(p,q,r)全体は三角形になります。

この図形上で微分幾何的な構造をいれて、
いろいろ調べるのが情報幾何の基本的な目標となります。

特に構造として加えるのは、以下のものです。

- リーマン計量
- 接続

今回は情報幾何で現れるリーマン計量と接続を定義し、
これらについて性質を述べようと思います。

最初に図形とその記号について定義しておきます。

$I$ を空でない有限集合とし, 位数を $n$ とする. この時,$I$ 上の測度全体が多様体となることを見る.

## 2.2 basic geometric objects

$\mathcal{F}(I) := \{f : I \to \mathbb{R}\}$ とすると,$\mathcal{F}(I)$ は $\mathbb{R}$ ベクトル空間になる. その基底として,

$$e_i(j) = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases}$$

が取れる. 実際 $f = \sum_{i=1}^{n} f^i e_i (f^i \in \mathbb{R})$ と一意にかけるので、確かにベクトル空間の基底となる. $F(I)^*$ を $F(I)$ から $\mathbb{R}$ への線形写像全体とする. また $F(I)^*$ を $S(I)$ と書く.

線形写像 $\sigma : \mathcal{F}(I) \to \mathbb{R}$ は $(I, P(I))$ の測度 $\mu$ を以下で与える.

$$\mu(J) := \sum_{j \in J} \sigma(e_j)$$

逆に $(I, P(I))$ 上の測度 $\mu$ は $\sigma(e_j) = \mu(\{j\})$ で $F^*(I)$ の元を定める. $\sigma$ 全体を $S(I)$ で表す.$S(I)$ は $n$ 次元ベクトル空間であり, 基底 $\delta^i$ を以下で定める.

$$\delta^i(e_j) = \begin{cases} 1 & i = j \\ 0 & otherwise \end{cases}$$

$S(I)$ は $\mathbb{R}$ ベクトル空間なので, 自然に多様体の構造が入る. また, 以降では以下が調べる基本的な対象である.

$$\begin{aligned} S_a(I) :=& \left\{\sum \mu_i \delta^i \mid \sum \mu_i = a\right\} \\ M(I) :=& \{\sum \mu_i \delta^i \mid \mu_i \geq 0\} \\ M_+(I) :=& \{\sum \mu_i \delta^i \mid \mu_i > 0\} \\ P(I) :=& \{\sum \mu_i \delta^i \mid \sum \mu_i = 1, \mu_i \geq 0\} \\ P_+(I) :=& \{\sum \mu_i \delta^i \mid \sum \mu_i = 1, \mu_i > 0\} \end{aligned}$$

$P_+(I) \subset M_+(I) \subset P(I)$ が部分多様体となっている.

今回の目標

統計上の道具を幾何的に定義する。

- Fisher計量の定義
- e-接続,m-接続の定義
- Markov Congruent Kernel で不変なテンソル場の話

Fisher計量 $M := M_+(I)$, $T_\mu M \cong S(I)$

$g_\mu : T_\mu M \times T_\mu M \to \mathbb{R}$

$$(f, g) \longmapsto \sum \frac{f_i g_i}{u_i}$$

をFisher計量という

$f: M_1 \to M_2$ が与えられた時

$$\frac{\partial}{\partial x_i} \longmapsto \sum \frac{\partial \psi_j \circ f \circ \varphi}{\partial x_i} \frac{\partial}{\partial y_j}$$

$$\begin{array}{ccc} U & \xrightarrow{\varphi} & M_1 \\ & & \downarrow f \\ V & \xrightarrow{\psi} & M_2 \end{array}$$

$$\frac{\partial \psi^{-1} \circ f \circ \varphi}{\partial x_i}$$

がinduceされる,
$v \in T_p M_1$ の行き先を
$\frac{\partial f}{\partial v}$ とかく

我々は $P_+(I)$ に興味があるので

$A, B \in T_\xi P_+(I)$ に対し, $p: P_+(I) \hookrightarrow M_+(I)$

$$\begin{aligned} g_\xi(A, B) &:= p^*(g)_\xi (A, B) \\ &= g_{p(\xi)}\left(\frac{\partial p}{\partial A}, \frac{\partial p}{\partial B}\right) \\ &= \sum_i \frac{1}{p_i(\xi)} \frac{\partial p_i}{\partial A} \cdot \frac{\partial p_i}{\partial B} \\ &= \sum p_i(\xi) \frac{\partial \log p_i}{\partial A} \cdot \frac{\partial \log p_i}{\partial B} \end{aligned}$$

で Fisher計量を定める。

e-connection, m-connection

$M := M_+(I)$, $g$: $M$上のRiemann metric

**Def** $M$: $C^\infty$-mfd, $E$: $M$上のbundle

$\nabla$が$E$上のconnectionとは
$\nabla: \Gamma(TM)\times\Gamma(E)\longrightarrow\Gamma(E)$
(これを$\nabla_X Y$で表す)

- $X$について$C^\infty$-linear
- $Y$についてlinear
- Leibniz rule
  $\nabla_X(fY)=(Xf)\cdot Y+f\nabla_X Y$

**Prop** $\nabla$: conn on $E$, $p\in M$
$X_1, X_2\in\Gamma(TM)$, $Y\in\Gamma(E)$
then $X_1(p)=X_2(p)\Rightarrow \nabla_{X_1}Y(p)=\nabla_{X_2}Y(p)$

$X_1, X_2$を$U\subset M$ open上に制限してもよい
($\nabla$は局所的に定まるので)
それは $f_1, f_2$ を1の分割で、
$f_1$を$V$上で1, $M\setminus U$上で0とすると、

$$\nabla_{(f_1+f_2)X_1}Y(p)=\underset{1}{\underset{\|}{f_1(p)}}\nabla_{X_1}Y(p)+\underset{0}{\underset{\|}{f_2(p)}}\nabla_{X_2}Y(p)$$

$=\nabla_{X_1}Y(p)$となる。
よって$X_1$を$U$上のベクトル場のときに示せば十分である。

同様に $Y = f_1 Y + f_2 Y$ とすると

$$\nabla_{X_1} f_1 Y(p) = \underbrace{X_1(f_1)}_{0}(p) + \nabla_X Y(p)$$
$$= \nabla_X Y(p).$$

$\nabla_{X_1} f_2 Y(p) = 0$ より, $Y$とU上と思て

この時、

$X_1 = \sum f_{1i} \frac{\partial}{\partial x_i}$, $X_2 = \sum f_{2i} \frac{\partial}{\partial x_i}$

$Y = \sum g_i \frac{\partial}{\partial x_j}$ とすると

$$\nabla_{X_1} Y = \sum_{i,j} \left( f_{1i} \nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} + \frac{\partial g}{\partial x_i}(p) \right)$$

$$\nabla_{X_2} Y = \sum_{i,j} \left( f_{2i} \nabla_{\frac{\partial}{\partial x_i}} \frac{\partial}{\partial x_j} + \frac{\partial g}{\partial x_i}(p) \right).$$

で $X_1(p) = X_2(p)$ より $f_{1i}(p) = f_{2i}(p)$
となるので一致する.

**Prop** $M: C^\infty$-mfd,
$\Phi: M \times \mathbb{R}^k \longrightarrow E$ と $x$
$s_i := \Phi(-, e_i) \in \Gamma$ ≡)

then $E$上の connection $\nabla$は

$$\nabla_{\square} s_i : \Gamma(TM) \longrightarrow \Gamma(E)$$

$(i=1,\cdots,k)$

一意に決まる

$Y=\sum f_i s_i$ とかけるので, trivial.

$$\left( \nabla_X Y = \nabla_X (\sum f_i s_i) = \sum (X(f_i) s_i + f_i \nabla_X s_i) \right.$$

<u>Def</u> $TM$上のconnection

$\nabla^{(m)}$, $\nabla^{(e)}$ を

$$\left( \begin{array}{l} TM = M \times S(I), \\ \partial_i : M \to TM \\ \quad u \mapsto (u, s^i) \end{array} \right. \text{とする.} \Big)$$

$$\tilde{\nabla}^{(m)}_A \partial_i = 0 \quad (\forall i=1,\cdots,m)$$

$$\tilde{\nabla}^{(e)}_A \partial_i = -\langle A, \partial_i \rangle \partial_i$$

この時, $B=\sum f_i \partial_i$ とかけるので

)より

この時，$B=\sum f_i\partial_i$ とおけるので．

$$\tilde{\nabla}^{(m)}_A B=\sum f_i \overset{0}{\overset{\shortparallel}{\nabla_A\partial_i}}+\sum A(f_i)\partial_i$$ となる

$$\tilde{\nabla}^{(e)}_A B=-\sum f_i\langle A,\partial_i\rangle\partial_i+\sum A(f_i)\partial_i$$
となる

曲線に沿うvector field $\dot{c}(t)$ に対し，
$V(t)=\nabla_{\dot{c}(t)}V$ となる

は linear，$\frac{D}{dt}(fV)=\frac{df}{dt}V+f\frac{D}{dt}V$

るものを共変微分という

$\nabla_{\dot{c}(t)}V=0$ となる時 X は$\nabla$についてCと並行

$c(0)=a$, $c(1)=b$ とした時
X を$\nabla^{(m)}$，Y を$\nabla^{(e)}$ について Cと並行とすると

$g_a(X,Y)=g_b(X,Y)$ となる

Amari-Chentsov TensorはFisher計量を３階のテンソル場にしたもの.
Fisher計量もAmari-Chetsov Tensorも以下のような性質で定義できる。

#### 2.3.1 Markov Kernel

**Definition 2.1.** $K : I \to P(J), i \mapsto K^i := \sum K^i_j \delta^j$ を *Markov Kernel* という. さらに、*Markov Kernel* が $A_i := \{j \in J \mid K^i_j\}$ とすると, これが $J$ の細分になる時,*congruent* という.

この時以下が成り立つ

**Theorem 2.2.** $P_+(I)$ 上の計量 $h^I$ が全ての *Markov congruent Kernel* に対し,

$$h^I_p(A, B) = h_{K_*(p)}(d_pK_*A, d_pK_*B) \tag{5}$$

となるのは定数倍を除き *Fisher Metric* のみである.

Amari Chentosv TensorについてもMarkov Congruent Kernel
について同様の性質で議論できる。

証明は例えば以下のように行われる
1. ベクトル場の取り方が変わらないことから重心上で
   Fisher Metricの定数倍で表せることを示す.
2. 重心と座標を入れ替えてた場合の変化で、
   定数倍が変わらないことを示す。

まとめ
ここまでで幾何的に出てきそうな道具を整備した。
- Fisher計量の定義
- 二種類の接続とその接続とFisher計量の関係
- 二種類の接続の差とAmari-Chentsov Tensorの関係
- Amari-Chentsov TensorとFisher計量の関係
    - 特にMarkov Congruent Kernelでの普遍性

今後具体的な統計的な対象を幾何的に翻訳できればと思う。