§14. 正準ダイバージェンス

統計多様体に対応するコントラスト関数は一意的ではない．双対平坦空間に対してはコントラスト関数の１つとして, 正準ダイバージェンスというものを考えることができる．
(M, g, ∇, ∇^*) を双対平坦空間とする．
まず, §13 において扱ったように局所的に考えよう．θ, η をそれぞれ ∇, ∇^* に関する M の近傍 U 上のアファイン座標系で,

$$\theta = (\theta^1, \theta^2, \dots, \theta^n),\ \partial_i = \frac{\partial}{\partial \theta^i}\ (i = 1, 2, \dots, n),$$

$$\eta = (\eta_1, \eta_2, \dots, \eta_n),\ \partial^j = \frac{\partial}{\partial \eta_j}\ (j = 1, 2, \dots, n)$$

と表しておき, θ と η は g に関して互いに双対的であるとする．また, ψ, φ を Legendre 変換

$$d\psi = \sum_{i=1}^n \eta_i d\theta^i,\ d\varphi = \sum_{i=1}^n \theta^i d\eta_i,\ \psi + \varphi = \sum_{i=1}^n \theta^i \eta_i \tag{$*$}$$

をあたえるポテンシャルとする．そこで, $p, q \in U$ に対して

$$\rho(p, q) = \psi(p) + \varphi(q) - \sum_{i=1}^n \theta^i(p)\eta_i(q)$$

とおく．ρ は g に関して互いに双対なアファイン座標系および Legendre 変換のポテンシャルの選び方に依存しないことが分かる．
このとき, $(*)$ の第３式より,

$$\rho(p, p) = 0.$$

また, $i = 1, 2, \dots, n$ とすると, $(*)$ の第１式より,

$$\begin{aligned}\rho(\partial_i | \cdot\,) &= \frac{\partial \psi}{\partial \theta^i} - \eta_i \\ &= 0.\end{aligned}$$

同様に, $(*)$ の第２式より,

$$\rho(\,\cdot\,|\partial_i) = 0.$$

更に, $j = 1, 2, \dots, n$ とすると,

$$\begin{aligned}\rho(\partial_i \partial_j | \cdot\,) &= \frac{\partial^2 \psi}{\partial \theta^i \partial \theta^j} \\ &= g_{ij}.\end{aligned}$$

よって, ρ は U 上の非退化計量 g をあたえるコントラスト関数である．ρ を正準ダイバージェンスといい,

$$\rho(p, q) = D(p \| q)$$

とも表す．
ここで, $k = 1, 2, \dots, n$ とすると,

$$\rho(\partial_i \partial_j | \partial_k) = 0.$$

一方, θ は ∇ に関するアファイン座標系であるから,

$$g(\nabla_{\partial_i}\partial_j, \partial_k) = 0.$$

したがって, 任意の $X, Y, Z \in \mathfrak{X}(U)$ に対して

$$g(\nabla_X Y, Z) = -\rho(XY|Z)$$

がなりたつ.
同様に,

$$g(Z, \nabla^*_X Y) = -\rho(Z|XY)$$

もなりたつから, ρ に対応する g に関して互いに双対なアファイン接続は ∇, ∇^* である.
特に, g が正定値の場合を考えよう.
$q \in U$ を固定しておき, $p \in U$ の関数 $f(p)$ を

$$f(p) = \sum_{i=1}^{n} \theta^i(p)\eta_i(q) - \psi(p)$$

により定める.
g の正定値性より, ψ は凸関数となり, f は $p = q$ において最大値 $\varphi(q)$ をとることが分かるから,

$$\varphi(q) = \max\left\{ \sum_{i=1}^{n} \theta^i(p)\eta_i(q) - \psi(p) \,\middle|\, p \in U \right\}$$

がなりたつ.
同様に,

$$\psi(p) = \max\left\{ \sum_{i=1}^{n} \theta^i(p)\eta_i(q) - \varphi(q) \,\middle|\, q \in U \right\}$$

がなりたつ.
よって,

$$\rho(p, q) \geq 0$$

で, $\rho(p, q) = 0$ となるのは $p = q$ のときに限る.
コントラスト関数は 2 点間の隔たりを表す距離のようなものであるが, 対称性や三角不等式といった距離の公理はみたさない. しかし, 正準ダイバージェンスに対しては次の Pythagoras の定理がなりたつ.

Pythagoras の定理　D を双対平坦空間 (M, g, ∇, ∇^*) 上の正準ダイバージェンス, p, q, r を互いに十分近い M 上の点とする. γ_1 を p と q を結ぶ ∇ に関する測地線, γ_2 を q と r を結ぶ ∇^* に関する測地線とする. γ_1 と γ_2 が q において g に関して直交するならば,

$$D(p\|r) = D(p\|q) + D(q\|r).$$

証明　g に関して互いに双対なアファイン座標系 θ, η および Legendre 変換のポテンシャル ψ, φ を上のように選んでおく.
アファイン座標系に関する Christoffel の記号はすべて 0 であるから, 測地線はアファイン座標系を用いると, 直線またはその一部として表すことができる.
よって, γ_1 は θ を用いて

$$\gamma_1(t) = t\theta(p) + (1 - t)\theta(q)$$

と表すことができる.
同様に, γ_2 は η を用いて

$$\gamma_2(t) = t\eta(r) + (1-t)\eta(q)$$

と表すことができる.
γ_1 と γ_2 は q において g に関して直交し, θ と η は g に関して互いに双対的だから,

$$\begin{aligned}
0 &= g(\gamma_1'(0), \gamma_2'(0)) \\
&= g\left(\sum_{i=1}^n (\theta^i(p) - \theta^i(q))\partial_i, \sum_{j=1}^n (\eta_j(r) - \eta_j(q))\partial^j\right) \\
&= \sum_{i=1}^n (\theta^i(p) - \theta^i(q))(\eta_i(r) - \eta_i(q)).
\end{aligned}$$

したがって,

$$\begin{aligned}
D(p\|q) + D(q\|r) - D(p\|r) &= \psi(p) + \varphi(q) - \sum_{i=1}^n \theta^i(p)\eta_i(q) + \psi(q) + \varphi(r) - \sum_{i=1}^n \theta^i(q)\eta_i(r) \\
&\quad - \left(\psi(p) + \varphi(r) - \sum_{i=1}^n \theta^i(p)\eta_i(r)\right) \\
&= \varphi(q) + \psi(q) - \sum_{i=1}^n \theta^i(p)\eta_i(q) - \sum_{i=1}^n \theta^i(q)\eta_i(r) + \sum_{i=1}^n \theta^i(p)\eta_i(r) \\
&= \sum_{i=1}^n \theta^i(q)\eta_i(q) - \sum_{i=1}^n \theta^i(p)\eta_i(q) - \sum_{i=1}^n \theta^i(q)\eta_i(r) + \sum_{i=1}^n \theta^i(p)\eta_i(r) \\
&= \sum_{i=1}^n (\theta^i(p) - \theta^i(q))(\eta_i(r) - \eta_i(q)) \\
&= 0.
\end{aligned}$$

すなわち,

$$D(p\|r) = D(p\|q) + D(q\|r).$$

□

N を C^∞ 級多様体, ∇ を N のアファイン接続, M を ∇ に関して自己平行な N の C^∞ 級部分多様体, ι を M から N への包含写像とする. このとき, M のアファイン接続 ∇^ι が定まる. §3 において扱ったことを思い出そう. また, ∇ が捩れをもたないならば, ∇^ι も捩れをもたず, ∇ の曲率が 0 ならば, ∇^ι の曲率も 0 となることが分かる.
特に, N に非退化計量 g があたえられ, (N, g, ∇, ∇^*) が双対平坦空間ならば, $(M, \iota^*g, \nabla^\iota, (\nabla^\iota)^*)$ も双対平坦空間となることが分かる.
更に, Pythagoras の定理より, 次がなりたつ.

定理 (N, g, ∇, ∇^*) を双対平坦空間とし, g は Riemann 計量で, D は N 上の正準ダイバージェンスとする. また, M を ∇^* に関して自己平行な N の C^∞ 級部分多様体とし, $p \in N$ とする. このとき, $q \in M$ が

$$D(p\|q) = \min\{D(p\|r) | r \in M\}$$

をみたすことと p と q を結ぶ ∇ に関する測地線が q において M と直交することとは同値.

関連事項 14. Ricci テンソル

アファイン接続の曲率を用いて, Ricci テンソルという $(0,2)$ 型のテンソル場を定めることができる.
M を C^∞ 級多様体, ∇ を M のアファイン接続, R を ∇ の曲率とする.
R は $(1,3)$ 型のテンソル場であるから, $X, Y \in \mathfrak{X}(M)$ を固定しておき, $Z \in \mathfrak{X}(M)$ とすると, Z から $R(Z,X)Y$ への対応は M の各点において接空間の線形変換を定める. この線形変換のトレースを $\mathrm{Ric}\,(X,Y)$ と表し, Ricci テンソルという. 定義より, Ricci テンソルは M 上の $(0,2)$ 型のテンソル場である.
(M,g) が Riemann 多様体で, ∇ が Levi-Civita 接続の場合を考えよう. このとき, Ricci テンソルは対称, すなわち

$$\mathrm{Ric}\,(X,Y) = \mathrm{Ric}\,(Y,X)$$

となる. 実際, 曲率の性質

$$R(X,Y)Z = -R(Y,X)Z$$

と Levi-Civita 接続の場合になりたつ性質

$$g(R(X,Y)Z,W) + g(Z,R(X,Y)W) = 0$$

および

$$g(R(X,Y)Z,W) = g(R(Z,W)X,Y)$$

を用いて計算すればよい. ただし, $X, Y, Z, W \in \mathfrak{X}(M)$ である. §2 において扱ったことを思い出そう.
また, 各 $p \in M$ に対して T_pM は内積空間となるが, 単位ベクトル $v \in T_pM$ に対して $\mathrm{Ric}\,(v,v)$ を v の Ricci 曲率という.
Ricci テンソルの対称性より, $u, v \in T_pM$ に対して

$$\mathrm{Ric}\,(u+v,u+v) = \mathrm{Ric}\,(u,u) + 2\mathrm{Ric}\,(u,v) + \mathrm{Ric}\,(v,v)$$

がなりたつから, 逆に, Ricci 曲率は Ricci テンソルを定めることも分かる.
Ricci 曲率は関連事項 13 において述べた断面曲率を用いて表すこともできる. $u, v \in T_pM$ が 1 次独立なとき, u, v の生成する T_pM の 2 次元部分空間に対する断面曲率を $K(u,v)$ と表すことにする. このとき, $\{v, e_2, e_3, \ldots, e_n\}$ が T_pM の正規直交基底ならば,

$$\mathrm{Ric}(v,v) = \sum_{i=2}^{n} K(v,e_i)$$

である.
更に, $\{e_1, e_2, \ldots, e_n\}$ が T_pM の正規直交基底のとき,

$$\sum_{i=1}^{n} \mathrm{Ric}\,(e_i,e_i)$$

を p におけるスカラー曲率という.
一般のアファイン接続に対しては Ricci テンソルは対称であるとは限らない. Ricci テンソルの対称性はアファイン接続の局所等積性と同値であることが分かる. ただし, M が n 次元 C^∞ 級多様体, ∇ が M のアファイン接続で, M の各点の近傍において $\nabla\omega = 0$ となる n 次微分形式 ω が存在するとき, ∇ は局所等積であるという.