数理解析研究所講究録
第682巻 1989年 81-105

曲指数分布族の逐次推定

東大 経済 岡本 一 (Ichi Okamoto)

1. はじめに

確率分布族は適当な条件のもとで特殊な多様体とみなせる. これを統計多様体と呼ぼう. 統計的推論における高次漸近理論では統計多様体の幾何構造が重要な役割を持つ. 曲指数分布族に関する非逐次推定法について考えよう. (標本の観測回数を前もって固定した場合の推定法を非逐次推定法と呼ぶ.) 曲指数分布族は自然に適当な指数分布族の部分多様体とみなされ、埋込曲率 (相対曲率, オイラー・スカウテン曲率) が定義される. このとき非逐次推定量の情報量損失は、この埋込曲率と深く関係することが知られている. (Amari, 1985)

逐次推定法はストッピングルールによって止められる観測 (これを逐次観測という) を伴った推定法である. Takeuchi and Akahira (1987) は1パラメータ分布族に関する逐次最尤推定法を提案した．ところがこの推定法の情報量損失は漸近的に0となる．これは逐次推定法を通して統計多様体の幾何構造に変化が生じることを意味している．1次元曲指数分布族に関して言えば，埋込曲率が消失することになる．

本論の目的は多次元曲指数分布族に関する逐次推定法を例にとり，高次漸近有効性にかかわる統計多様体の共形幾何構造を考察することである．

まず統計多様体の共形変換を定義する．そして逐次観測を通して与えられる，統計多様体の微分同型写像が共形変換であることを示す．次に曲指数分布族に関する逐次推定法を定義し，その推定法に関する情報量損失及び二次漸近分散を求める．このとき，情報量損失が共形埋込曲率と呼ばれる量と関係することがわかる．最後に逐次推定法が，漸近分散の意味で2次有効となる条件を与える．

付録では，1次漸近分散についての分散安定化を考察する．Shiryayev and Greenwood (1987) は1次自己回帰過程に対する分散安定化逐次推定法を提案した．ここでは逐次推定によって分散安定化を可能とする条件を与える．

ここでリーマン幾何学における共形変換について紹介しておこう．リーマン多様体は、可微分多様体Sと計量テンソル$g_{ij}(\theta)$によって与えられる．そして2つのリーマン多様体(S,g)と$('S,'g)$の間の微分同型写像が共形変換であるとは、

$$'g_{ij}=\rho\cdot g_{ij}$$

となる，滑らかな正値スカラ$\rho(\theta)$が存在することを言う．つまりSに共形変換を与えると長さの尺度が変化する．しかもその変化の仕方が点に応じて異なるのである．物理学的に言えば，物体に熱源を近づけ，等方的に伸縮させることにあたる．従って，物体の幾何構造は変化する．とくに接続や曲率は特殊な変化をする．逆に言うと，適当な共形変換によって曲率や接続の値をある程度制御することが可能になる．

ここでは，この共形変換の概念を統計多様体の幾何学，すなわち情報幾何学に応用する．

2．統計多様体における共形変換

まず 共通な σ 有限測度μに関する確率密度函数の空間Sを

$$S=\{p(x,\theta)\ ;\ \theta=(\theta^1,\theta^2,\cdots,\theta^n)\in \Theta\}$$

とおく．ここでθは$p(x,\theta)$を特定するためのn次元実ベクト

ルパラメータである．さらに S について次の条件が成立すると仮定しよう．

(A.1) 全ての $p(x,\theta)$ は x の函数として共通の台を持つ．

(A.2) 各 $p(x,\theta)$ は θ の函数として適当な階数まで微分可能である．

(A.3) $l(x,\theta)=\log p(x,\theta)$ とおく．各点 θ において

$$\partial_i l(x,\theta) \qquad i=1,2,\cdots,n$$

は線型独立である．但，∂_i は $\partial/\partial\theta^i$ を意味する．

(A.4) $\{\partial_i l(x,\theta)\}$ のモーメントは適当な次数まで存在する．

(A.5) 以下で扱う任意の函数 $f(x,\theta)$ について

$$\partial_i \int f(x,\theta)\,p(x,\theta)\,d\mu = \int \partial_i [f(x,\theta)p(x,\theta)]\,d\mu$$

が成立する．

(A.6) Θ は n 次元ユークリッド空間 R^n と同相である．

このとき，S について計量 g_{ij} 及び対称テンソル T_{ijk} がそれぞれ

$$g_{ij}(\theta)=E_\theta[\partial_i l(x,\theta)\,\partial_j l(x,\theta)],$$

$$T_{ijk}(\theta)=E_\theta[\partial_i l(x,\theta)\,\partial_j l(x,\theta)\,\partial_k l(x,\theta)]$$

と定義される．そこで，これから S を統計多様体 (S,g,T) とみなすことにする．

S の α-接続 $\Gamma^{(\alpha)}_{ijk}$ は，

$$\Gamma^{(\alpha)}_{ijk}=E_\theta[\partial_i\partial_j l(x,\theta)\,\partial_k l(x,\theta)]+\frac{1-\alpha}{2}T_{ijk}$$

と与えられる．とくに$(\alpha=1)$－接続は指数接続と呼ばれ，この指数接続に関する幾何量は添字に(e)を付けることで表される．また$(\alpha=-1)$－接続は混合接続と呼ばれ，混合接続に関する幾何量は添字に(m)を付けることにより示される．

Sのα－リーマン・クリストフェル曲率は

$$R^{(\alpha)}_{ijk}{}^{m}=\partial_i\Gamma^{(\alpha)}_{jk}{}^{m}-\partial_j\Gamma^{(\alpha)}_{ik}{}^{m}+\Gamma^{(\alpha)}_{ir}{}^{m}\Gamma^{(\alpha)}_{jk}{}^{r}-\Gamma^{(\alpha)}_{jr}{}^{m}\Gamma^{(\alpha)}_{ik}{}^{r}$$

と定義される．但 $\Gamma^{(\alpha)}_{ij}{}^{m}=\Gamma^{(\alpha)}_{ijk}\,g^{km}$ である．またα－リッチテンソル及びα－曲率スカラは、それぞれ

$$R^{(\alpha)}_{ij}=R^{(\alpha)}_{kij}{}^{k}$$
$$R^{(\alpha)}=R^{(\alpha)}_{ij}\,g^{ij}$$

と与えられる．（この論文ではアインシュタインの規約を用いる．従ってA_iB^iなどの量は$\sum_{i=1}^{n}A_iB^i$を意味するものとする．）

次に統計多様体の共形変換を定義しよう．

定義 2.1

統計多様体(S,g,T)と$(\tilde{S},\tilde{g},\tilde{T})$の間の微分同型$\xi$が共形変換であるとは，

$$\tilde{g}_{ij}=\rho\cdot g_{ij}$$
$$\tilde{T}_{ijk}=\rho\,(T_{ijk}+3g_{(ij}\,r_{k)})$$

を満たす，滑らかな正値スカラ$\rho(\theta)$が存在することをいう．

但，$r_i=\partial_i \log\rho$，$3g_{(ij}r_{k)}=g_{ij}r_k+g_{jk}r_i+g_{ki}r_j$ である．このとき，微分幾何学の習慣に倣い $\tilde{S}$ の計量及び対称テンソル N を

$$'g_{ij}=\frac{1}{\rho}\tilde{g}_{ij}=g_{ij}$$
$$'T_{ijk}=\frac{1}{\rho}\tilde{T}_{ijk}=T_{ijk}+3g_{(ij}r_{k)}$$

と定義し直す．

さて，

$$x=(x_1,x_2,\cdots,x_N)$$

を互いに独立に同一分布 $p(x,\theta)\in S$ に従う，標本列としよう．但，観測回数 N は次の条件を満たす逐次観測方式により決定されるものとする．

(B.1) $(n+1)$ 個目の標本を観測するかどうかは，それまでの n 個の標本に従って決定する．

(B.2) $E_\theta[N]=K\nu(\theta)$ ，$Var_\theta[N]=O(K)$

$E_\theta[N^i]=O(K^i)$ $\quad i=1,2,3,4$

ここで K は十分大なる正数，$\nu(\theta)$ は滑らかな正値スカラとする．

このとき x の同時確率密度函数は，

$$p(x,\theta)=\prod_{i=1}^{N}p(x_i,\theta)$$

で与えられる．そこで分布族 $\tilde{S}$ を

$$'S = \{ p(x,\theta) = \prod_{i=1}^{N} p(x_i,\theta) \ ; \theta \in \Theta \}$$

とおく．$'S$もまた統計多様体とみなせるので幾何量を計算しよう．その際，次の補題が有用である．

補題2.2 (Takeuchi and Akahira, 1987)

$l(x,\theta) = \log p(x,\theta)$とする．このとき$\theta$に関して滑らかな函数$Y(x,\theta)$について

$$E_\theta[Y(x,\theta)\partial_i l(x,\theta)] = \partial_i E_\theta[Y] - E_\theta[\partial_i Y]$$

が成立する．但，$E_\theta[Y]$はθについて微分可能であると仮定する．

実際に計量$\tilde{g}$及び対称テンソル$\tilde{T}$を計算してみると

$$\tilde{g}_{ij} = K\nu g_{ij}$$
$$\tilde{T}_{ijk} = K\nu(T_{ijk} + 3g_{(ij}s_{k)})$$

となる．但，$s_k = \partial_k \log \nu$とおく．これは逐次観測によって$S$が$'S$に共形変換されたことを意味する．そこで$'S$の計量と対称テンソルを改めて

$$'g_{ij} = g_{ij}$$
$$'T_{ijk} = T_{ijk} + 3g_{(ij}s_{k)}$$

と定義しておく．共形変換後のα-接続も

$$'\Gamma^{(\alpha)}_{ijk} = \Gamma^{(\alpha)}_{ijk} + \frac{3(1-\alpha)}{2} g_{(ij}s_{k)} - g_{ij}s_k$$

となる．

3．曲指数分布族の幾何学

以下ではSを指数分布族と考え，

$$S=\{p(x,\theta)=\exp(\theta^i x_i-\psi(\theta))\ ;\ \theta\in\Theta\cong R^n\}$$

とおく．Sの幾何量は

$$g_{ij}(\theta)=\partial_i\partial_j\psi(\theta)$$

$$T_{ijk}(\theta)=\partial_i\partial_j\partial_k\psi(\theta)$$

で与えられる．またS上には自然な座標系(θ^i)の他に，座標変換$\eta_i=\partial_i\psi(\theta)$によって定まる座標系$(\eta_i)$が存在する．この$(\eta_i)$を期待値座標系と呼ぶ．

次に曲指数分布族Mを

$$M=\{p(x,\theta(u))=\exp(\theta^i(u)\cdot x_i-\psi(\theta(u)))\ ;\ u\in U\cong R^m\}$$

とおく．このとき$(u^a)_{a=1,\cdots,m}$はMの座標系となる．さらに$\theta=\theta(u)$はUからΘへの滑らかな単射であると仮定しよう．するとMはSの部分多様体となるので，MをSの中へ写像$\theta=\theta(u)$により埋込むことができる．

さてSからMへの射影$\hat{u}$が与えられたと仮定する．このとき補助多様体$A(u)$は，

$$A(u)=\{\eta\in S\ ;\ \hat{u}(\eta)=u\}$$

と定義される．$A(u)$の座標系を$(v^{\kappa})_{\kappa=m+1,\cdots,n}$とおく．但，簡単のため，$v=0$で$A(u)$は$M$と交わっているものとする．すると$S$上には（正確にいうと$S$における$M$のある管状近傍には）混合座標系$(u^a, v^{\kappa})$が定まる．以下では添字$\alpha,\beta,\gamma,\cdots$は$(w^{\alpha})$-座標系に関する量を，添字$a,b,c,\cdots$は$(u^a)$-座標系に関する量を，また添字$\kappa,\lambda,\mu,\cdots$は$(v^{\kappa})$-座標系に関する量をそれぞれ表すものとする．

(θ^i)-座標系から(w^{α})-座標系へのヤコビアンを$B_{\alpha}^{i}=\partial_{\alpha}\theta^i$とおくと，$M$の計量及び対称テンソルは

$$g_{\alpha\beta}(u,0)=B_{\alpha}^{i}B_{\beta}^{j}g_{ij}(u,0)$$

$$T_{\alpha\beta\gamma}(u,0)=B_{\alpha}^{i}B_{\beta}^{j}B_{\gamma}^{k}T_{ijk}(u,0)$$

となる．また(η_i)-座標系から(w^{α})-座標系へのヤコビアンを$B_{\alpha i}=\partial_{\alpha}\eta_i$とおくと，$M$の$\alpha$-接続は

$$\Gamma^{(\alpha)}_{\beta\gamma\delta}(u,0)=\frac{1}{2}(1-\alpha)T_{\beta\gamma\delta}+(\partial_{\beta}B_{\gamma}^{i})B_{\delta i}$$

となる．とくに任意の$u\in M$に対して$g_{a\kappa}(u,0)=0$（つまり$M\perp A(u)$）ならば，MのSに対するα-埋込曲率と$A(u)$のSに対するα-埋込曲率がそれぞれ次のように定義される．

$$H^{(\alpha)}_{ab\kappa}=\Gamma^{(\alpha)}_{ab\kappa},$$

$$H^{(\alpha)}_{\kappa\lambda a}=\Gamma^{(\alpha)}_{\kappa\lambda a}.$$

5章以降，Mに関する推定量はSからMへの射影であるとみなされる．そこで射影をその補助多様体の幾何構造に従い

分類する．まずSからMへの滑らかな射影全体のクラスを$P_M(S)$とおく．

__定義3.1__

(1) $P_M(S)$に属する射影が一致射影であるとは，任意の点$u \in M$に対して　$u \in A(u)$が成立することをいう．

(2) 一致射影$\hat{u}$に関して，M上で$g_{a\kappa} \equiv 0$が成立するとき$\hat{u}$は直交射影と呼ばれる．

(3) 直交射影$\hat{u}$に関して，M上で$\overset{(m)}{H_{\kappa\lambda c}} \equiv 0$が成立するとき，$\hat{u}$は混合横断射影と呼ばれる．

4．曲指数分布族の共形変換

Sに対して，2章で定義した逐次観測を適用しよう．するとSと同時に部分多様体であるMも共形変換される．その像を$'M$とする．$'M$の幾何量は，

$$'g_{\alpha\beta} = g_{\alpha\beta}$$
$$'T_{\alpha\beta\gamma} = T_{\alpha\beta\gamma} + 3g_{(\alpha\beta}A_{\gamma)}$$

となる．とくに$'M$の混合接続は，

$$'\Gamma^{(m)}_{abc} = \Gamma^{(m)}_{abc} + 2g_{c(a}A_{b)}$$

で与えられる．但，$2g_{\alpha(\beta}A_{\gamma)} = g_{\alpha\beta}A_\gamma + g_{\alpha\gamma}A_\beta$である．

$M \perp A(u)$ となる場合，共形変換後も直交関係は保存されるので，とくに $'M$ の指数埋込曲率は，次式で定義される．

$$'H^{(e)}_{ab\kappa} \equiv {'\Gamma}^{(e)}_{ab\kappa} = H_{ab\kappa} - g_{ab} \Delta_\kappa .$$

また $A(u)$ の混合埋込曲率は

$$'H^{(m)}_{\kappa\lambda a} = H^{(m)}_{\kappa\lambda a} - 2g_{a(\kappa}\Delta_{\lambda)} = H^{(m)}_{\kappa\lambda a}$$

となるので，共形変換に関して不変である．

そこで 曲指数分布族Mにおける新たな幾何量を定義しよう．

定義4.1

(1) S における M の α－平均曲率ベクトルは，

$$H^{(\alpha)}_\kappa = \frac{1}{m} H^{(\alpha)}_{ab\kappa}\, g^{ab}$$

で定義される．また M の α－平均曲率スカラは

$$\| H^{(\alpha)}_M \| = \sqrt{H^{(\alpha)}_\kappa H^{(\alpha)}_\lambda g^{\kappa\lambda}}$$

で定義される．

(2) S における M の α－共形埋込曲率は次式で定義される．

$$\overline{H}^{(\alpha)}_{ab\kappa} = H^{(\alpha)}_{ab\kappa} - g_{ab} H^{(\alpha)}_\kappa$$

この α－共形埋込曲率は S に対する任意の共形変換に対して不変（つまり共形不変量）である．

5．標本平均にもとづく逐次推定法

次のような問題を考えよう。

$$\mathbf{x} = (x_1, x_2, \cdots, x_N)$$

は互いに独立に同一分布 $p(x, \theta(u)) \in M$ に従う標本列とする。ただし N はさきにあげた条件 (B.1) 及び次の条件 (B.2') を満たす逐次観測によって決まるものと仮定する。

(B.2') $E_u[N] = K\nu(u, 0)$, $Var_u[N] = O(K)$

$E_u[N^i] = O(K^i)$ $(i = 1, 2, 3, 4)$

ここで K が十分大なる正数とする。

ここで $\mathbf{x}$ から真のパラメータ u を推定したい。情報量損失あるいは二次漸近分散の意味で最適な推定法はどんなものか?

条件 (B.1), (B.2') を満たす逐次観測を伴った推定法のクラスを $\tilde{\Sigma}_M$ とおく。しかし我々はさらに条件を加える。

(B.3) 観測回数 N は標本平均 $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ に関する次式:

$$N = K\nu(\bar{x}) + c(\bar{x}) + \varepsilon. \qquad \cdots (5.1)$$

を満たす。但, ε は $O_p(1)$ であり $\bar{x}$ と漸近的独立, ならびに $E[\varepsilon] = o(1)$ を満たすと仮定する。

そこで, 推定量 $\hat{u}$ は $\hat{u} = \hat{u}(\bar{x})$ とする。

(注) 標本平均 $\bar{x}$ は $(\eta_i) = (\bar{x}_i)$ とおくことにより S 上の1点とみなすことができる。従って滑らかな正値スカラ $\nu(\eta)$ $(\eta \in S)$ に対して $\nu(\bar{x})$ を定義することができる。$c(\cdot)$ につい

ても同様である．そして推定量はこの場合，SからMへの射影にほかならない．

条件(B.3)は，xの代わりに$\bar{x}$を利用することを意味している．ところで，統計量Yの情報量損失は，

$$\Delta g_{ab}(Y) = E[\mathrm{Cov}[\partial_a l(x, \theta(u)), \partial_b l(x, \theta(u)) \mid Y]]$$

で与えられる．式(5.1)を満たす$\bar{x}$の情報量損失を直接計算すると，次の定理を得る．

定理5.1

(5.1)式を満たす標本平均は漸近的に2次十分統計量である．すなわち

$$\Delta g_{ab}(\bar{x}) = \frac{\mathrm{Var}(\varepsilon)}{K v} g_{ab} + O(K^{-2}) = O(K^{-1})$$

従って　2次漸近分散，あるいは推定法の情報量損失の意味では，xの代わりに$\bar{x}$を用いることが正当化される．

そこで条件(B.1)，(B.2)，(B.3)を満たす推定法のクラスをΣ_Mとおく。我々に残された自由度はスカラー$v(u,v)$ $(v \neq 0)$の値，及び推定量$\hat{u}$のとり方のみである．しかし1次漸近有効性に関しては，$\hat{u}$のとり方のみが関係する．

定理5.2

(1) Σ_M に属する推定法が一致推定法（$\hat{u} \xrightarrow{i.p.} u$ $(K\to\infty)$）である必要十分条件は，$\hat{u}$が一致射影であることである.

(2) Σ_M に属する一致推定法が一次有効である必要十分条件は，$\hat{u}$が直交射影であることである. 但，一次有効推定法とは，その推定量$\hat{u}$について，$K\to\infty$のとき

$$\sqrt{K\nu}\,(\hat{u}-u) \sim N(0,\, g^{ab})$$

となるものをいう.

証明は，非逐次推定の場合と同様なので省略する.

6. 情報量損失と漸近分散

推定法πの情報量損失は，πの推定量$\hat{u}$の情報量損失として定義される. そこで，直接計算することにより次の定理が得られる.

定理6.1

Σ_Mに属する推定法の情報量損失は，その推定量$\hat{u}$が直交射影ならば，

$$\Delta g_{ab} = ({}'H_M^{(e)})^2_{ab} + \frac{1}{2}(H_A^{(m)})^2_{ab} \qquad \cdots (6.1)$$

で与えられる．但，

$$({}'H_M^{(e)})^2_{ab} = {}'H^{(e)}_{ac\kappa}\,{}'H^{(e)}_{bd\lambda}\,g^{cd}g^{\kappa\lambda} = (H^{(e)}_{ac\kappa} - g_{ac}A_\kappa)(H^{(e)}_{bd\lambda} - g_{bd}A_\lambda)\,g^{cd}g^{\kappa\lambda}$$

$$(H_A^{(\cdot)})^2_{ab} = H^{(m)}_{\kappa\lambda a}\,H^{(m)}_{\mu\sigma b}\,g^{\kappa\mu}g^{\lambda\sigma}$$

である．

式(6.1)の右辺のうち，第一項はS上のスカランに関する項であり，$\hat{u}$にはよらない．また第二項は$\hat{u}$のみに依存し，逐次観測とは無関係である．従って次の事実が成立する．

<u>定理6.2</u>

Σ_Mに属する推定法の情報量損失が最小（正定値行列の意味で）となるためには，次の2条件がともに成立することが必要十分である．

(1) 推定量$\hat{u}$は混合横断射影である．

(2) スカラン$\nu(u,v)$について$A_\kappa = H_\kappa^{(e)}$が成立する．

このとき情報量損失の最小値は

$$\Delta g_{ab} = (\bar{H}_M^{(e)})^2_{ab}$$

となる．但，

$$(\bar{H}_M^{(e)})^2_{ab} = \bar{H}^{(e)}_{ac\kappa}\,\bar{H}^{(e)}_{bd\lambda}\,g^{cd}g^{\kappa\lambda}$$

である．

<u>証明</u>

式(6.1)の2つの項はともに正定値であることを注意しておく．まず(1)のもとでは

$$\Delta g_{ab} = ({}'H^{(e)}_{u})^2_{ab} \qquad \cdots (6.2)$$

だから，(2)のもとで(6.2)が最小をとることを示す．簡単のためトレースで比較しよう．(一般に2階共変テンソル A_{ab} のトレースは $A_{ab}g^{ab}$ で与えられる．)

(6.2)から

$$\mathrm{tr}\,({}'H^{(e)}_{u})^2_{ab} = (H^{(e)}_{u})^2_{ab}g^{ab} + m\{(s_\kappa - H^{(e)}_\kappa)(s_\lambda - H^{(e)}_\lambda) - H^{(e)}_\kappa H^{(e)}_\lambda\}g^{\kappa\lambda}$$

$$\geq \mathrm{tr}(H^{(e)}_{u})^2_{ab} - m\|H^{(e)}_{u}\|^2 \qquad \cdots (6.3)$$

が得られる．(6.3)式の等号成立条件は $s_\kappa = H^{(e)}_\kappa$ である．そしてこのとき，

$${}'H^{(e)}_{ab\kappa} = H^{(e)}_{ab\kappa} - g_{ab}H^{(e)}_\kappa = \overline{H}^{(e)}_{ab\kappa}$$

が成立する．(証おわり)

推定法 π に関する<u>不足数</u> $\Delta(\pi)$ を

$$\Delta(\pi) = \frac{1}{m}\Delta g_{ab}(\pi)\cdot g^{ab}$$

として定義する．いま，$s_\kappa = H^{(e)}_\kappa$ とする1次有効推定法を π_1 とおき，$s_\kappa = 0$ とする1次有効推定法を π_2 とおく．(推定量については同一の推定量 $\hat{u}$ をとっておく．) このとき，(6.3)式から，

$$\Delta(\Pi_1) = \Delta(\Pi_2) - \| H_M^{(e)} \|^2$$

が得られる．つまり Π_1 は Π_2 より，$\|H_M^{(e)}\|^2$ だけ不足数の意味で有利だと言える．

次に Σ_M に属する推定法に関する漸近分散を求めよう．しかし，非逐次推定の場合と同様に推定量のバイアスを補正する必要がある．

<u>定理6.3（岡本，1988a）</u>

Σ_M に属する推定法に関する，漸近的バイアスは，

$$E[\sqrt{K\nu}(\hat{u}^{*a} - u^a)] = -\frac{1}{2\sqrt{K\nu}} \Gamma_{\alpha\beta}^{(m)a} g^{\alpha\beta} + O(K^{-1})$$

となる．但，$\hat{u}$ は直交射影である．

従って，バイアス補正推定量 $\hat{u}^*$ を

$$\hat{u}^{*a} = \hat{u}^a + \frac{1}{2K\nu(\hat{u},0)} \Gamma_{\alpha\beta}^{(m)a}(\hat{u}) g^{\alpha\beta}(\hat{u})$$

と定義すると

$$E[\sqrt{K\nu}(\hat{u}^{*a} - u^a)] = O(K^{-1})$$

となる．

バイアス補正された，直交射影 $\hat{u}^*$ に関しては，

$$E[K\nu(\hat{u}^{*a} - u^a)(\hat{u}^{*b} - u^b)] = g^{ab} + \frac{1}{K\nu}\left\{\frac{1}{2}({}'\Gamma_M^{(m)})^{2ab} + \Delta g^{ab}(\hat{u})\right\} + O(K^{-2})$$

が成立する。但，

$$({}'\Gamma_M^{(m)})^2_{ab} = {}'\Gamma^{(m)}_{cda}\ {}'\Gamma^{(m)}_{efb}\ g^{ce} g^{df}$$

$$({}'\Gamma_M^{(m)})^{2ab} = ({}'\Gamma_M^{(m)})^2_{cd}\ g^{ca} g^{db}$$

$$\Delta g^{ab} = \Delta g_{cd} \cdot g^{ca} g^{db}$$

である．そこで次のように漸近分散が得られる．

<u>定理 6.4 (岡本, 1988a)</u>

Σ_Mに属するバイアス補正1次有効推定法の2次漸近分散は，

$$E[K\nu \cdot (\hat{u}^{*a} - u^a)(\hat{u}^{*b} - u^b)] = g^{ab} + \frac{1}{K\nu}\left\{\frac{1}{2}({}'\Gamma_M^{(m)})^{2ab} + ({}'H_M^{(e)})^{2ab} + \frac{1}{2}(H_A^{(m)})^{2ab}\right\} + O(K^{-2})$$

となる．ここで

$$({}'H_M^{(e)})^{2ab} = ({}'H_M^{(e)})^2_{cd}\ g^{ca} g^{db}$$

$$(H_A^{(m)})^{2ab} = (H_A^{(m)})^2_{cd}\ g^{ca} g^{db}$$

である．とくに最小2次漸近分散は，

$$E[K\nu \cdot (\hat{u}^{*a} - u^a)(\hat{u}^{*b} - u^b)] = g^{ab} + \frac{1}{K\nu}\left\{\frac{1}{2}({}'\Gamma_M^{(m)})^{2ab} + (\overline{H}_M^{(e)})^{2ab}\right\} + O(K^{-2})$$

で与えられる．但，

$$(\overline{H}_M^{(e)})^{2ab} = (\overline{H}_M^{(e)})^2_{cd}\ g^{ca} g^{db}.$$

2次漸近分散を最小とする推定法を2次有効推定法と呼ば

う．結局，次の2条件：

(1) $A_\kappa = H_\kappa^{(e)}$ ；すなわち $\bar{x}=\eta(\hat{u},\hat{v})$ とおいたとき，

$$N = K\nu(\hat{u},0)\{1+H_\kappa^{(e)}(\hat{u},0)\hat{v}^\kappa\}+O_p(1) \qquad \cdots (6.4)$$

(2) $\hat{u}$ は，混合~~横断射影~~であり，バイアスは

$$E[\hat{u}] = u + O(K^{-\frac{3}{2}})$$

となる．

を満たす推定法が2次有効となることがわかった．Takeuchi and Akahira (1987) の逐次最尤推定法も，上の意味で，2次有効であることがわかっている．そこで，その推定法を多次元に拡張してみよう．つまり 次の条件：

(I) 観測回数 N は

$$-\frac{1}{m}\sum_{i=1}^{N}\partial_a\partial_b l(x,\hat{u})g^{ab}(\hat{u}) = K\nu(\hat{u},0)+C(\hat{u},0)+\varepsilon$$

を満たす．

(II) 推定量 $\hat{u}$ はバイアス補正最尤推定量である．

を満たすものを考える．このとき，やはりこの推定法が2次有効であることがわかる．

(注)．スカラ $C(\eta)$ は，

$$C(\eta) = -\frac{1}{2}(\partial_\beta A_\gamma - \Gamma_{\beta\gamma}^{(m)\alpha}A_\alpha - A_\beta A_\gamma)g^{\beta\gamma}$$

で与えられる．また，(B.1)，(B.2)，(B.3)を満たすストッピング・ルールを実際に求めることは，困難だが可能である．(Akahira and Takeuchi, 1988)

7. おわりに

以上，逐次推定法に関する漸近有効性について論じてきた．ここでは，もう少し直観的に逐次推定法，あるいは逐次観測について考えよう．

非逐次推定法に関して情報量損失とMの埋込曲率が関係する理由は岡本(1988b)が簡単に述べた．$A_x=0$ とおく逐次推定法についても全く同様である．つまり $\bar{x}$ が曲率方向におちても，反対方向におちても観測回数Nは殆ど変化しない．この結果，埋込曲率に関する項が，そのまま情報量損失に残る．

一方，$A_x=H_x^{(e)}$ とおく逐次推定法では，$\bar{x}$ が曲率方向におちたとき，Nは相対的に大きくとられる．Nが大きくなれば情報量も当然大きくなる．つまり情報量損失について悪影響を及ぼす方向で，情報量を補償しようとする働きが，この逐次観測には存在する．結果として，埋込曲率のうち，平均曲率分が，情報量損失において回復される．

物理的類推をするのも興味深い．かりにMが熱によって変形する物体だと考えよう．すると観測回数N（あるいは $\nu(u,v)$）は，熱に対応している．従って，標語的には，「逐次観測は，統計多様体のある部分を温め，他の部分を冷やすことにより変形を加える機構である．とくに $A_x=H_x^{(e)}$ となる逐次観測は，

曲指数分布族Mの曲率方向を温め，反対方向を冷やす．そして埋込曲率を減らす働きがある.」 と言える.

付録. 分散安定化について

ここでは，$\nu(u,0)$の値をuによって変えることにより，漸近分散の第一項を操作することについて述べる.

定義A.1

曲指数分布族M上にスカラ$\nu(u,0)$が存在して，νg_{ab}が点uによらない定数行列になるとき，Mは分散安定化可能という.

分散安定化という語の由来は，次の事実に因る.

定理A.2

曲指数分布族Mが分散安定化可能であるならば，

$$\sqrt{K}(\hat{u}-u) \sim N(0, C^{ab}) \qquad (K\to\infty)$$

となる逐次推定法が存在する．但，ここでC^{ab}は点uによらない定数行列（正確には2階反変テンソル）である.

証明は，定理5.2(2)から明らかである．とくに $C^{ab}=(\nu g_{ab})^{-1}$ となる．

さて，微分幾何学でよく知られているように，νg_{ab} を定数とするための必要十分条件は，M が(0)-共形平坦，つまり，M とユークリッド空間 R^m の間に共形変換が存在することである．(丹野，1976) 具体的には次の定理が成立する．

定理A.3

(1) $\dim M \equiv m = 1$ の場合，M は任意の座標系に関して分散安定化が可能である．

(2) $m=2$ の場合，M は適当な座標系 (u^a) において分散安定化可能である．

(3) $m=3$ の場合，M が適当な座標系 (u^a) において分散安定化可能となる必要十分条件は，

$$\overset{(0)}{C}_{abc} \equiv (\overset{(0)}{\nabla}_b \overset{(0)}{R}_{ca} - \overset{(0)}{\nabla}_c \overset{(0)}{R}_{ba}) - \frac{1}{4}(g_{ab}\partial_c \overset{(0)}{R} - g_{ac}\partial_b \overset{(0)}{R})$$

が，M 上各点で 0 となることである．但，

$$\overset{(0)}{\nabla}_b \overset{(0)}{R}_{ca} = \partial_b \overset{(0)}{R}_{ca} - \overset{(0)}{\Gamma}{}_{bc}{}^{d} \overset{(0)}{R}_{da} - \overset{(0)}{\Gamma}{}_{ba}{}^{d} \overset{(0)}{R}_{cd} .$$

(4) $m \geqq 4$ の場合，M が適当な座標系 (u^a) において分散安定化となるための必要十分条件は，

$$\overset{(0)}{C}{}_{abc}{}^{d} \equiv \overset{(0)}{R}{}_{abc}{}^{d} - \frac{1}{m-2}(\delta_b{}^d \overset{(0)}{R}_{ca} - \delta_c{}^d \overset{(0)}{R}_{ba} + \overset{(0)}{R}{}_b{}^d g_{ca} - \overset{(0)}{R}{}_c{}^d g_{ba})$$
$$+ \frac{1}{(m-1)(m-2)} \overset{(0)}{R} (\delta_b{}^d g_{ca} - \delta_c{}^d g_{ba})$$

がM上各点で0となることである．但，$R_a^{(0)b} = R_{ac}^{(0)} g^{cb}$であり，また$\delta_a^b$はクロネッカーのデルタを表す．

この定理における，$C_{abc}^{(0)}$，$C_{abc}^{(0)}{}^d$は，(0)－ワイル共形曲率と呼ばれる．

例えば，$N(\mu, \sigma^2)$における，平均μと分散σ^2の同時逐次推定を考えよう．このとき

$$g_{ab}(\mu, \sigma^2) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}$$

となるので，$\nu(\mu, \sigma^2) = \sigma^2$とすれば，分散安定化が可能である．実は$N(\mu, \sigma^2)$は定曲率空間であり（Amari, 1985）従って(0)－共形平坦であることがわかっている．

ところで，Mが分散安定化可能でない場合でも，次の事実は必ず成立する．

<u>定理A.4</u>

(1) 正値スカラ$\nu_1(u, 0)$をM上で適当にとれば，$\nu_1 g_{ab}$の最小固有値を点によらない定数とすることができる．（このことを擬似分散安定化という．）

(2) 正値スカラ$\nu_2(u, 0)$をM上で適当にとれば，$\nu_2 g_{ab}$の

行列式を貞によらない定数とすることができる.(これを一般化分散安定化と呼ぶ.)

(3) 正値スカラ $\nu_3(u,0)$ を適当にとることによって $\nu_3 g_{ab} f^{ab}$ (f^{ab}は,任意の正定値行列)を貞によらない定数とすることができる.

例えば,(3)により　十分小なる正数εに対して

$$E[K\cdot\Sigma_{a=1}^{m}(\hat{u}^a-u^a)^2]=\varepsilon+O(K^{-1})$$

となる,逐次推定法が存在することがわかる.

これらの分散安定化は,一般の正則な分布族に関する推定問題や時系列問題にも簡単に拡張できる.また,ここでは逐次推定にもとづく分散安定化について述べたが,非逐次推定に関する分散安定化についてはYoshizawa (1971), Amari (1985)を見よ.

参考文献

Akahira, M. and Takeuchi, K. (1988). Third order asymptotic efficiency of the sequential maximum likelihood estimation procedure, submitted for the publication.

Amari, S. (1985). Differential-Geometrical Methods in Statistics, Lecture Note in Statistics 28, Springer Verlag, Berlin.

Shiryayev, A.N. and Greenwood, P.E. (1987). A minimax property of sequential estimates in autoregressive models, (submitted for the publication).

Takeuchi, K. and Akahira, M. (1987). Second order asymptotic efficiency in terms of asymptotic variances of the sequential maximum likelihood estimation procedure, Statistical theory and Data analysis II (ed. by Matsushita, K.), North-Holland, Amsterdam.

Yoshizawa, T. (1971). A geometry of parameter space and its statistical interpretation. Memo TYH-2, Harvard Univ.

岡本 一. (1988 a). 逐次推定の微分幾何学, 東京大学 大学院 工学系研究科 情報工学専門課程修士論文.

岡本 一. (1988 b). 消えた曲率の謎, 数理科学 9月号, 26-31, サイエンス社.

丹野修吉 (1976) 多様体の微分幾何学, 192-194, 実教出版.