

漸近不偏推定量の幾何的構成

間野 修平 数理・推論研究系 教授

0 本ポスターの内容

母数の関数の推定に興味があるとする。期待値が真の値に一致する推定量を不偏推定量という。その導出は一般には難しいので、漸近的に不偏な推定量を求めることを考える。この種の問題へのベイズ的接近は古い(恐らく Hartigan 1964 に始まる)。微分方程式を解くことに帰着し、母数が多次元の場合は偏微分方程式が現れるが、その一般的扱いが困難なためか、一貫した枠組みが無かった。一方、偏微分方程式の解析における微分幾何の有効性は古くから認識されている。本研究では、問題を微分幾何を用いて定式化することで、漸近不偏推定量をベイズ的に構成する枠組みを与えた。本発表は廣瀬雅代氏(九州大学)との共同研究である。

1 モデル多様体の脱規格化

母数 $\xi \in \Xi$, 確率変数 $x \in X$ に対し, 未知の確率測度 $p(x)$ と非負有限測度 $q(x; \xi)$ の Kullback-Leibler (KL) divergence は

$$D\{p||q(\cdot; \xi)\} = \int_X q(x; \xi) dx - 1 + \int_X p(x) \log p(x) dx - \int_X p(x) \log q(x; \xi) dx,$$

で与えられる。確率測度のモデル $q_0(\cdot; \xi)$ により $q(\cdot; \xi) = q_0(\cdot; \xi)z(\xi)$, $z(\xi) := \int_X q(x; \xi) dx > 0$ と表す。最尤推定は標本 (x_1, \dots, x_n) の経験分布と $q_0(\cdot; \xi)$ の KL-divergence の最小化であるが, 事前分布の対数を $\tilde{l}(\xi) = n\{1 - z(\xi) + \log z(\xi)\} \leq 0$ とする MAP(maximum a posteriori) 推定は経験分布と脱規格化されたモデル $q(\cdot; \xi)$ の KL-divergence の最小化である。 $q_0(\cdot; \xi)$ 上で不偏になるように事前分布 $e^{\tilde{l}(\xi)}$, つまり規格化定数 $z(\xi)$ を定めたい。

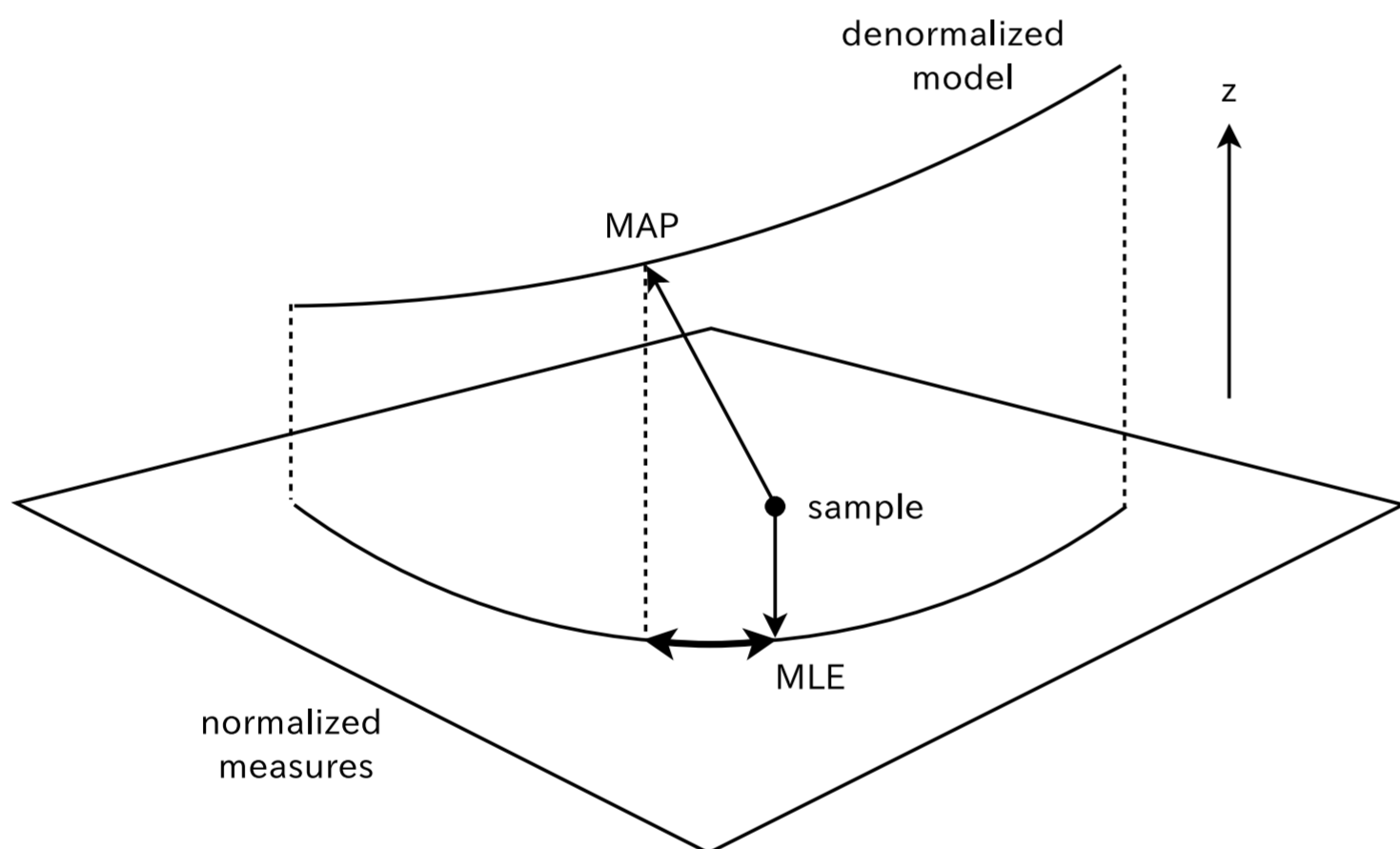


図1 モデル多様体の脱規格化

2 α -Laplace作用素と漸近不偏性

α 接続 $\nabla^{(\alpha)}$, ただし $\Gamma_{ijk}^{(\alpha)} := \mathbb{E}[(\partial_i \partial_j l)(\partial_k l)] + (1 - \alpha)S_{ijk}/2$, $\alpha \in \mathbb{R}$ に対し α -Laplace作用素を $\Delta^{(\alpha)} f := \nabla^{(\alpha)i} \nabla_i^{(\alpha)} f$ で定義する。 l は尤度の対数, $S_{ijk} := \mathbb{E}[(\partial_i l)(\partial_j l)(\partial_k l)]$, $\partial_i = \partial/\partial \xi^i$ である。 Fisher 計量テンソルを $g_{ij} = \mathbb{E}[(\partial_i l)(\partial_j l)]$, ベクトル場 X, Y の内積を $\langle X, Y \rangle = g_{ij} X^i Y^j$ で表す。

定理 1 (HM). 正則条件の下, 不偏推定可能函数 f と事前分布の対数 $\tilde{l}(\xi)$ が

$$\Delta^{(-1)} f + 2\langle \partial \tilde{l}, \partial f \rangle = o(n^{-1}) \quad (1)$$

を満たすとき, $\hat{\xi}$ を ξ の MAP 推定量として, $f(\hat{\xi})$ は $f(\xi)$ の 2 次漸近不偏推定量(偏りが $o(n^{-1})$)であり, 完備十分統計量があれば $f(\hat{\xi})$ は一様最小分散不偏推定量(UMVUE)に $O(n^{-1})$ まで一致する。

注意 1. $f(\hat{\xi})$ は最尤推定量と同じく 2 次漸近有効である。

Hirose, Lahiri (2018) の混合効果モデルに関する結果は次の例である。

系 1. 1 次元モデル多様体を考える。正則条件の下, 不偏推定可能函数 $f(\xi)$, $f'(\xi) > 0$ に対し, 事前分布の対数 $\tilde{l}(\xi)$ が

$$e^{\tilde{l}(\xi)} \propto \frac{\{g(\xi)\}^{1/4}}{\sqrt{f'(\xi)}} e^{\frac{1}{4} \int^\xi S_1(\tilde{\xi}) d\tilde{\xi}}$$

を満たせば, $\hat{\xi}$ を ξ の MAP 推定量として, $f(\hat{\xi})$ は $f(\xi)$ の 2 次漸近不偏推定量で, 完備十分統計量があれば UMVUE に $O(n^{-1})$ まで一致する。

d 次元多様体においてアフィン接続 ∇ が各点で $\nabla \omega = 0$ を満たす d 次微分形式 ω をもつとき ∇ を局所等積といい, ω を平行体積要素という。次の命題が成り立つ。

命題 1 (Takeuchi-Amari 2005). モデル多様体が 0 でない α_0 について α_0 平坦のとき, 任意の α について α 接続は局所等積である。

α_0 平坦多様体における α 接続に対する平行体積要素の密度 $g^{1-\frac{\alpha}{2\alpha_0}}$ を α_0 平坦多様体における α 平行事前分布とよぶ。 0 平行事前分布は α_0 によらず Jeffreys 事前分布である。

系 2. 特に, モデル多様体が 0 でない α について平坦であれば, α アフィン座標 ξ について $\hat{\xi}$ を $(\alpha - 1)/2$ 平行事前分布による MAP 推定量として, $\Delta^{(\alpha)} f = 0$ を満たす函数 $f(\xi)$ の推定量 $f(\hat{\xi})$ は 2 次漸近不偏で, 完備十分統計量があれば UMVUE に $O(n^{-1})$ まで一致する。

多様体 M 上の r 次元の分布 \mathcal{D} とは, M の各点 p において接空間 $T_p M$ の C^∞ 級 r 次元部分空間 \mathcal{D}_p を対応させるもので, 各点近傍にベクトル場 X_1, \dots, X_r が存在し, 近傍の任意の点 q において \mathcal{D}_q の基底をなすものをいう。 \mathcal{D} に属する任意のベクトル場 X, Y について $[X, Y]$ が \mathcal{D} に属するとき, \mathcal{D} は包摂的であるという。

定理 2 (Frobenius). C^∞ 級多様体上の分布 \mathcal{D} が完全可積分であるための必要十分条件は \mathcal{D} が包摂的であることである。

命題 2 (HM). 正則条件の下, 不偏推定可能函数 f に対し, 事前分布の対数を

$$\tilde{l}(t) = -\frac{1}{2} \int^t \frac{\Delta^{(-1)} f}{\langle \partial f, \partial f \rangle} d\tilde{t} + \text{const.}, \quad (2)$$

$t = f(\xi)$ で与えれば, $\hat{\xi}$ を ξ の MAP 推定量として, $f(\hat{\xi})$ は $f(\xi)$ の 2 次漸近不偏推定量で, 完備十分統計量があれば UMVUE に $O(n^{-1})$ まで一致する。

$\tilde{l}(\xi)$ についての 1 階偏微分方程式 (1) が $(d - 1)$ 次元の包摂的分布を定めるとき, Frobenius の定理より, (2) はレベル集合 $N_{t_0} = \{\xi : f(\xi) = t_0\}$, $t_0 \in f(M)$ からなる M の余次元 1 の葉層 $\{N_t : t \in f(M)\}$ を与える。

3 例

2 変量正規分布 $N_d(\xi, I)$ における平均ベクトルの原点からの測地距離 r の平方の函数の推定に興味があるとする。モデル多様体は Euclid 空間 \mathbb{R}^2 , 原点を中心とする円の族 $\{N_r : r > 0\}$, $N_{r_0} = \{(r, \omega) : r = r_0, \omega \in S^1\}$ はその葉層をなす。測地距離 r の推定を考える。(2) より $\tilde{l}(r^2) = -\log r/2$, 規格化定数は $r > 1$ のとき $z(r) = 1 + \sqrt{\log r/n} + O(n^{-1})$ である。十分統計量は標本平均 \bar{x} , r の 2 次漸近不偏で UMVUE に $O(n^{-1})$ まで一致する推定量は $r(\hat{\xi}) = |\bar{x}| \{1 - 1/(2n\bar{x}^2)\} + o_p(n^{-1})$ である。図 2 に 2 次漸近不偏性を与える規格化定数を示す。簡単のため経験分布はモデル多様体の上にあるとした。 $r(\hat{\xi})$ は最尤推定量 $|\bar{x}|$ を原点に向けて縮小したものである。

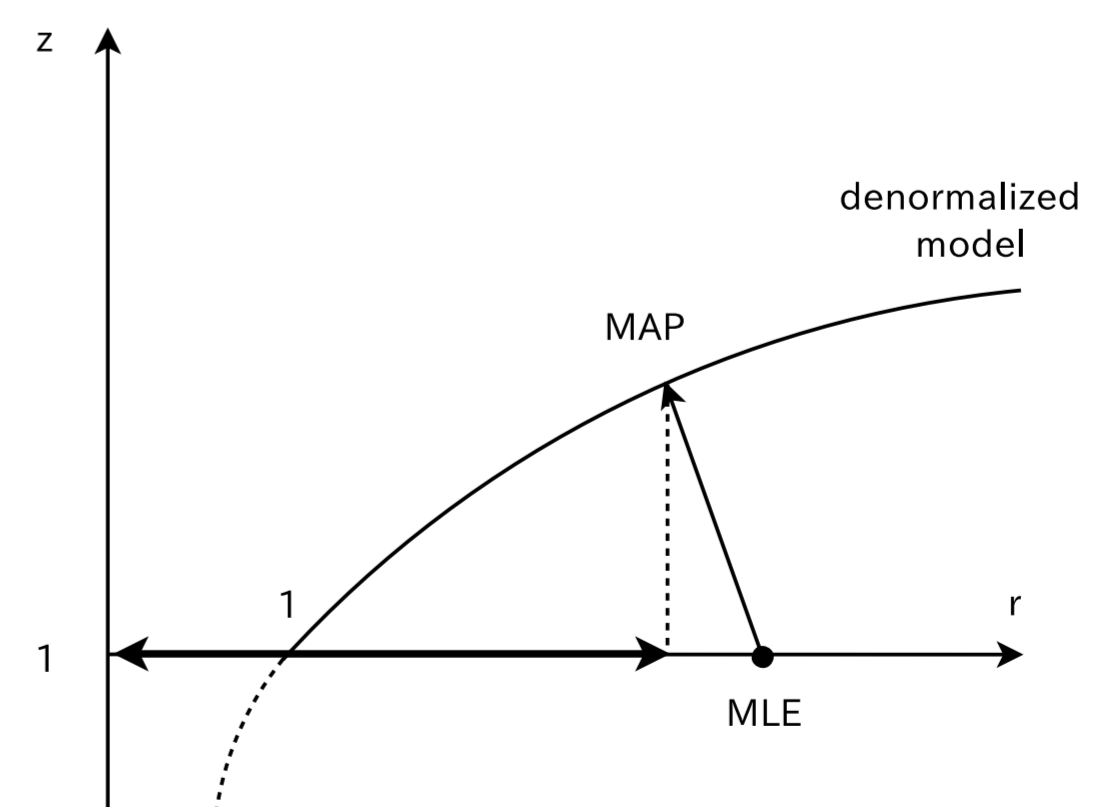


図2 2次漸近不偏性を与える規格化定数。

参考文献

A Bayesian construction of asymptotically unbiased estimators. arXiv: 2011.14747