2024 年 5 月 24 日 統計数理研究所 オープンハウス

統計多様体上の測地距離の高次漸近有効不偏推定

間野 修平　統計基盤数理研究系 教授

1　本ポスターの概要

パラメトリックモデルに基づく推論において，同じ確率分布の族に属する2つの確率分布の間の距離を測ることを考えよう．そのような距離は色々考えられるが，パラメタ空間の各点が確率分布に対応するから，パラメタ空間における2点間の距離を測っても良いだろう．では，どのような距離が妥当だろうか．

例えば，正規分布$\mathrm{N}(\mu,\sigma^2)$は平均μと標準偏差σで定まるから，2次元のパラメタ空間に値をとる確率分布の族をなすが，$\mathrm{N}(0,1)$と$\mathrm{N}(\mu,\sigma^2)$の距離として三平方の定理より従うユークリッド距離

$$\sqrt{|\mu-0|^2+|\sigma-1|^2} \tag{1}$$

を使うのは妥当とは言い難い．なぜなら，パラメタ空間を平坦(ユークリッド空間)とする根拠がないからである．

確率分布の属を，パラメタを局所座標系，フィッシャー情報行列をリーマン計量とする微分可能多様体とみなしたものを統計多様体と呼ぶ．微分多様体の2点間の距離としては測地距離が基本的である．2点を結ぶ曲線のうち長さが最小になるものを測地線といい，測地距離はその長さである．ユークリッド空間においては，測地線は直線，測地距離はユークリッド距離である．統計多様体の2点間の距離として測地距離を用いるのは自然だろう．

ただし，統計的推論においては，データが統計多様体上の未知の点(上の例では(μ,σ))に対応する確率分布に従って生成されたと仮定し，データからその点と所与の点(上の例では$(0,1)$) の測地距離を推定する必要がある．測地距離の表現は複雑になりがちで，推定が難しそうに見えるかもしれない(式(3)を見よ)．

発表者と廣瀬雅代氏(九州大学)は，パラメトリックモデルに基づく推論において，2次漸近有効不偏推定量，すなわち，標本サイズをnとして誤差を$o(n^{-1})$に抑えて，不偏（期待値が真の値に等しい）かつ有効（不偏推定量の分散の下限を満たす）な推定量を求める一般的な枠組みを与えた(Hirose+M, arXiv:2011.14747)．本ポスターは，論文に記載した統計多様体上の測地距離の推定への応用の紹介である．

2　結果

最尤推定量の期待値を真の値の周りで$O(n^{-1})$まで漸近展開することで次の結果を得た．

定理 1 (HM20+)**.** パラメタを$\xi\in\Xi$とする．正則条件の下，対数尤度lに罰則項として函数$\tilde{l}:\Xi\to\mathbb{R}$を加えた罰則付き最尤推定量$\hat{\xi}$と函数$f:\Xi\to\mathbb{R}$に関して，次が成り立つ．

(i) $f(\hat{\xi})$は$f(\xi)$の2次漸近有効推定量である．

(ii) 特に$\tilde{l}$が偏微分方程式

$$2\langle \mathrm{grad}\tilde{l}, \mathrm{grad} f\rangle + \Delta f = 0 \tag{2}$$

を満たすならば，$f(\hat{\xi})$は$f(\xi)$の2次漸近不偏推定量でもある．ただし

$$\Gamma_{ij,k} := \mathbf{E}[(\partial_i\partial_j l)(\partial_k l)] + \mathbf{E}[(\partial_i l)(\partial_j l)(\partial_k l)], \quad \Delta f := \nabla^i\nabla_i f.$$

注意 1. 最尤推定量$f(\hat{\xi}_{\mathrm{MLE}})$には一般に$O(n^{-1})$ の偏りがある．定理1は適切な罰則項を付けることで最尤推定量を2次漸近有効性を維持しつつ2次漸近不偏に改善できることを主張する．

式(2)は準線形1階偏微分方程式だから一般に解くことができて(解の存在と一意性を示せる)，幾何学的な意味を持つ．すなわち，$\tilde{l}$の水準を座標とする次元を追加し，積分曲面$(\phi=\phi_0)$の$\tilde{l}$の水準による葉層を統計多様体Mに射影するとMのfの水準による葉層を与えるように，積分曲面が埋め込まれる(図1)．fは積分曲面に接するモンジュ軸を定める．

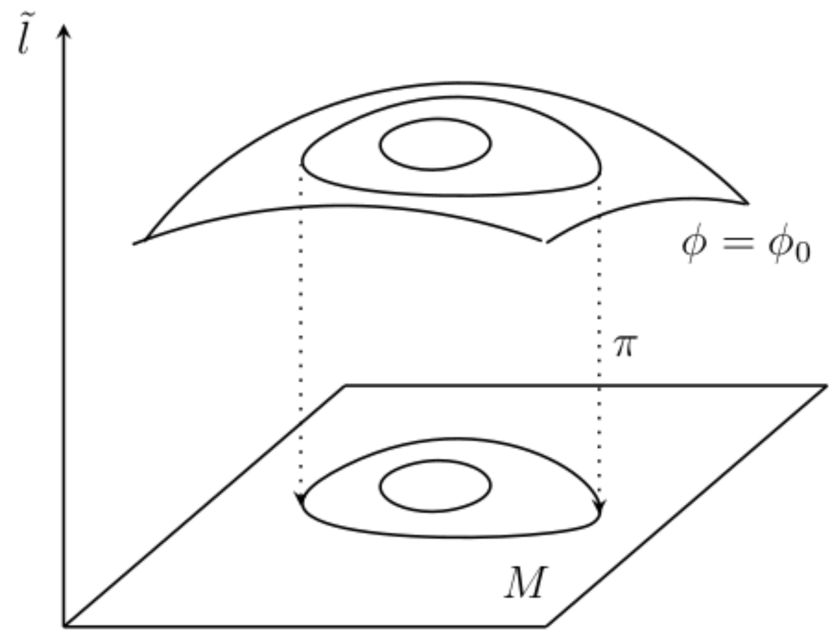

図 1: 積分曲面の埋め込みと葉層

測地距離の2乗に関して次の系が従う．測地距離の任意の函数についても同様であるが，以下，表現が簡潔な測地距離の2乗について記す．

系 1. 統計多様体上の1点ζを固定し，点ξとの間の測地距離を$t(\xi)$で表す．以下の罰則項の下，$t^2(\hat{\xi})$は$t^2(\xi)$の2次漸近有効不偏推定量である．

$$\tilde{l}(\xi) = -\frac{1}{2}\log\left\{t^d\sqrt{h}\right\} + \frac{1}{4}\int^t \mathbf{E}[(\partial_i l)(\partial_j l)(\partial_k l)]g^{jk}\xi^i du.$$

ただし，h_{ij}はζにおける正規座標系のフィッシャー計量である．

3　例：位置尺度族

原点対称な確率密度$p_0(z)$, $z\in\mathbb{R}$の位置尺度族とは，密度が

$$p(x)dx = \frac{1}{\sigma}p_0\left(\frac{x-\mu}{\sigma}\right)dx$$

で与えられる確率分布の族をいう．$l_0(z)=\log p_0(z)$, $\mathbf{E}(l_0'')=-R^2$として，フィッシャー計量は$g_{ij}=(R/\sigma)^2\delta_{ij}$であるが，これは2次元双曲空間の上半平面モデルのポアンカレ計量である．点$(0,1)$と点$\xi=(\mu,\sigma)$を結ぶ測地線(円弧)を図2に示す．その測地距離は

$$t(\xi) = R\left[\log\left\{\nu+\sqrt{\nu^2-1}\right\}\right], \quad \nu := \frac{\mu^2+\sigma^2+1}{2\sigma} \tag{3}$$

であり，$t^2(\hat{\xi})$を2次漸近有効不偏推定量とする$\hat{\xi}$を与える罰則項は

$$\tilde{l}(\xi) = -\log t(\xi) - \frac{1}{4}\log h(\xi) + \frac{c}{4R^2}\log\sigma$$

である．cはl_0から定まる定数である．$h(\xi)$の表現は省略する．積分曲面及び上半平面の葉層を図3，数値実験($n=100$)の結果を表に示す．

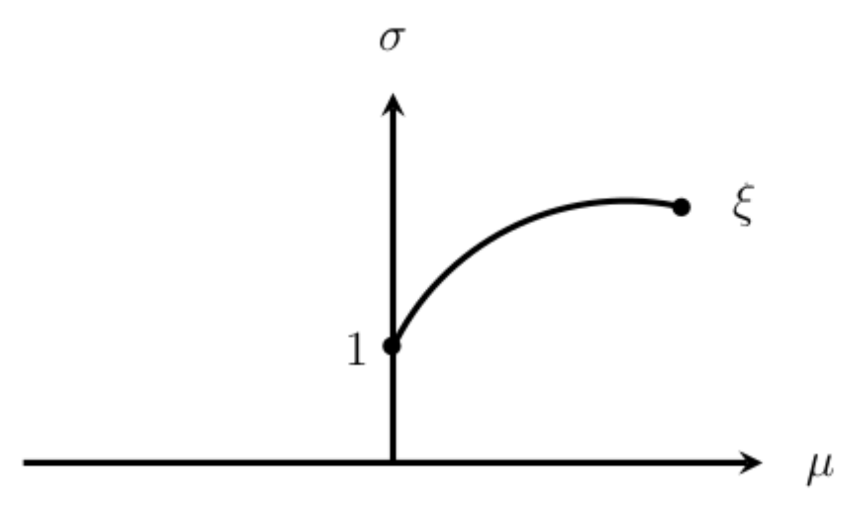

図 2: 測地線

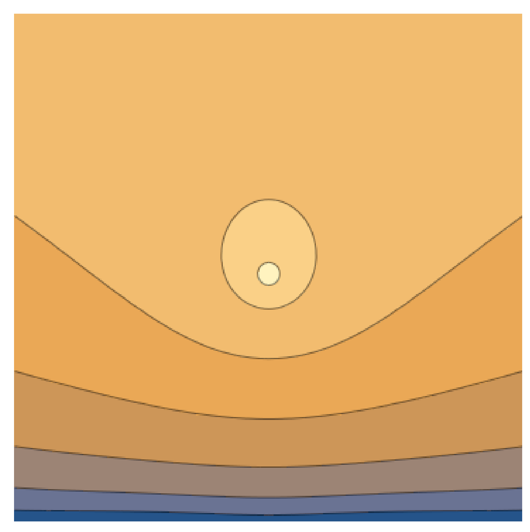

図 3: 図1を真上から見る

	$(\mu,\sigma)=(1,1)$		$(\mu,\sigma)=(5,1)$		$(\mu,\sigma)=(0,0.1)$	
	$t^2=1.85252$		$t^2=21.70696$		$t^2=10.60380$	
	偏り	平均2乗誤差	偏り	平均2乗誤差	偏り	平均2乗誤差
最尤法	+0.03610	0.07872	+0.13328	0.91014	+0.10132	0.44693
本方法	−0.00058	0.07568	−0.02003	0.88393	−0.00263	0.43295
	$(\mu,\sigma)=(5,0.1)$		$(\mu,\sigma)=(0,5)$		$(\mu,\sigma)=(5,5)$	
	$t^2=61.85059$		$t^2=5.18058$		$t^2=10.69656$	
	偏り	平均2乗誤差	偏り	平均2乗誤差	偏り	平均2乗誤差
最尤法	+0.22886	2.57769	−0.01906	0.20499	+0.03608	0.43819
本方法	−0.05227	2.51639	+0.00126	0.20603	−0.00145	0.42903