カイ二乗ダイバージェンスに基く 無情報事前分布の情報幾何

大阪大学・基礎工　田中 冬彦

客観ベイズ分析において, Bernardo らが提案している reference prior [1] は, ジェフリーズ事前分布の導出など一定数の成果をあげてきた．その定義は以下のような汎関数の最大化問題の解で与えられる．今, 統計モデル $\{p(x|\theta) : \theta \in \Theta \subset \mathbf{R}^k\}$ が与えられた時, 事前分布 $\pi \in \mathcal{P}(\Theta)$ について事後分布から事前分布への KL ダイバージェンス $D(\pi(\theta|x), \pi(\theta))(=: D_\pi(x))$ の期待値を

$$J(\pi) := E^\theta E^X [D_\pi(X)] = \iint D_\pi(x) p(x|\theta) \mathrm{d}x \pi(\theta) \mathrm{d}\theta$$

とおく．Reference prior [1] は $J(\pi)$ を最大にする事前分布として定義され, 特に $n \to \infty$ ではジェフリーズ事前分布 π_{J} によって最大になる．Reference prior は多くの拡張が考えられてきたが, 最近, Liu *et al.* [2] は, $D(p, q)$ を他の多くのダイバージェンスにおきかえても π_{J} が $J(\pi)$ の主要項を最大にすることを示した．ところが, $D(p, q)$ としてカイ二乗ダイバージェンスをとると $\int E[\pi(\theta|X)|\theta]\mathrm{d}\theta$ の最大化に帰着し, π_{J} とは異なる事前分布が出てくることが示された.

本発表では, Liu *et al.* [2] の結果を幾何学的な形に書き変え, そこから新たな結果が得られることを示す．g_{ij} を Fisher 情報行列, g^{ij} をその逆行列, $g := \det(g_{ij})$ とし $T_{jkm} = E[\partial_j l \partial_k l \partial_m l]$, $T_j := T_{jkm} g^{km}$ とする．このとき, $\phi := \pi/\pi_{\mathrm{J}}$, $\pi_{\mathrm{J}} := \sqrt{g}$ とおいて, Liu *et al.* [2] の結果を ϕ の汎関数として書き直すと

$$(4\pi)^{\frac{k}{2}} n^{-\frac{k}{2}} \int E[\pi(\theta|X)|\ \theta]\mathrm{d}\theta = \int \left[1 + \frac{1}{n} \left\{ -\frac{1}{4} \left\| \mathrm{d}\log\phi + \frac{T}{4} \right\|^2 + r(\theta) \right\} + o(n^{-1}) \right] \sqrt{g(\theta)} \mathrm{d}\theta$$

となる．ここで $\|A\|^2 := A_i A_j g^{ij}$, $r(\theta)$ は ϕ に依存しない項．さて, $\mathrm{d}\log\phi = -\frac{1}{4}T$ を満たす非負のスカラー関数 ϕ が存在する時, $\pi_{\chi^2} \propto \phi \cdot \pi_{\mathrm{J}}$ で事前分布 π_{χ^2} を定義しよう．このとき, 以下が成立.

定理 1: π_{χ^2} はカイ二乗ダイバージェンスによる $J(\pi)$ を $n \to \infty$ で最大にする.

π_{χ^2} の存在の必要十分条件は可積分条件 $\partial_j T_i - \partial_i T_j = 0$ である．これは Takeuchi and Amari [3] による α-parallel prior の存在条件でもあり, 両者の関連性が予想される．実際, この予想は正しい.

定理 2: α-parallel prior $\pi^{(\alpha)}$ の満たす方程式は, $\phi^{(\alpha)} := \pi^{(\alpha)}/\pi_{\mathrm{J}}$ とおくと $\mathrm{d}\log\phi^{(\alpha)} = -\frac{\alpha}{2}T$. 従って, π_{χ^2} は $\frac{1}{2}$-parallel prior である.

Takeuchi and Amari [3] では指数型分布族で α-parallel prior を与えていた ($\alpha = 1/2$ で Liu *et al.* [2] の例と一致)．これは γ-平坦 ($\gamma \neq 0$) な統計モデルに容易に拡張でき, $\pi^{(\alpha)}(\eta) \propto (\pi_{\mathrm{J}}(\eta))^{1-\frac{\alpha}{\gamma}}$. ($\{\eta\}$ はアファイン座標系.) とくに $\alpha = 1/2$ とすると $\pi_{\chi^2} \propto (\pi_{\mathrm{J}}(\eta))^{1-\frac{1}{2\gamma}}$.

REFERENCES

[1] J. M. Bernardo: Reference posterior distributions for Bayesian inference. *J. R. Statist. Soc. B*, **41** (1979), 113–147.

[2] R. Liu, A. Chakrabarti, T. Samanta, J. K. Ghosh and M. Ghosh: On divergence measures leading to Jeffreys and other reference priors. *Bayesian Analysis*, **9** (2014), 331–370.

[3] J. Takeuchi and S. Amari: α-parallel prior and its properties. *IEEE Trans. Info. Theory*, **51 no.3** (2005), 1011–1023.