局所情報による統計的推論

金森 敬文 (名古屋大)

概要

離散サンプル空間上の確率分布の推定について考察する．この問題では統計モデルの規格化定数の計算が困難であり，これを回避するために，擬似尤度や合成尤度など，規格化定数を必要としない損失を用いる方法が提案されている．これらの推定量では，サンプル空間における近傍の情報のみを用いて推定量を構成する．本稿ではこのような局所情報を用いる推定量の漸近分散と近傍系との関連について考察する [1].

1. ランダム近傍系から定義される Z-推定量

離散サンプル空間 $\mathcal{X}$ 上の統計モデルを $p_\theta(x) = \widetilde{p}_\theta(x)/Z_\theta, \theta \in \Theta \subset \mathbb{R}^d$ とし，適当な正則条件を仮定しておく．実用上よく用いられる統計モデルでは，規格化定数 $Z_\theta = \sum_{x\in\mathcal{X}} \widetilde{p}_\theta(x)$ の計算が困難なことが多い．この困難を回避するために，規格化定数を必要としない推定量がいくつか提案されている．例えば擬似尤度 (pseudo-likelihood) や，その一般化である合成尤度 (composite likelihood) などがある．これらの推定量の推定関数は，ランダム化した近傍系に関する期待値として与えられる場合がある．本稿では，ランダム化近傍を許容する推定量を確率的局所 Z-推定量とよぶ．そして局所情報のみを用いる推定量を，確率的局所 Z-推定量のランダム近傍上での期待値として構成する．

定義 1 (ランダム近傍系・確率的局所 Z-推定量・局所 Z-推定量 [1])**.** $\mathcal{X}$ の各点 $x \in \mathcal{X}$ に，x を含む集合の族 (近傍系) $\mathcal{N}_x \subset 2^{\mathcal{N}}$ とその上の条件付き確率 $q(C|x), C \in \mathcal{N}_x$ が付随しているとき，$\{\mathcal{N}_x, q(\cdot|x)\}_{x\in\mathcal{X}}$ をランダム近傍系という．サンプル点 $x \in \mathcal{X}$ と $x \in C \in \mathcal{N}_x$ に対して定義される関数 $\widetilde{f}(x, C) \in \mathbb{R}^d$ で任意の $C \in \mathcal{N}_x$ に対して $\mathbb{E}_{\theta,q}[\widetilde{f}|C] = \sum_x p_{\theta,q}(x|C)\widetilde{f}(x, C) = 0$ を満たすものを推定関数とする Z-推定量を，確率的局所 Z-推定量という ($p_{\theta,q}(x|C)$ は $p_\theta(x)q(C|x)$ から定義される条件付き分布)．さらに，条件付き分布 $q(C|x)$ による確率的局所 Z-推定量 $\widetilde{f}(x, C)$ の期待値 $f_\theta(x) = \mathbb{E}[\widetilde{f}|x] = \sum_{C\in\mathcal{N}_x} \widetilde{f}(x, C)q(C|x)$ を推定関数とする Z-推定量を，局所 Z-推定量という．

合成尤度のあるクラスや擬似尤度の推定関数は局所 Z-推定量である．実際，$\mathcal{X} = \{0, 1\}^d$ 上の確率分布の擬似尤度は次のように表せる：$\mathcal{N}_x = \{\{x, x_{-k}\} : k = 1, \ldots, d\}$ (x_{-k} は $x \in \{0, 1\}^d$ の第 k 要素を反転した点)，$q(C|x) = 1/d, C \in \mathcal{N}_x, \widetilde{f}(x, C) = \nabla \log(\widetilde{p}_\theta(x)/\sum_{z\in C} \widetilde{p}_\theta(z))$.

2. 近傍系と漸近分散

関数 $g(x, C) \in \mathbb{R}^d$ は $(x, C) \in \mathcal{X} \times \mathcal{N}_x$ に対して定まり，$\mathbb{E}_{\theta,q}[g] = \sum_{x,C} g(x, C)p_\theta(x)q(C|x) = 0$ を満たすとする．このような関数の集合は，適当な正則条件のもとで，分布 $p_\theta(x)q(C|x)$ における接空間をなす．関数 $g(x, C)$ を $\mathbb{E}[g|C]$ に変換することで確率的局所 Z-推定量が得られ，さらに $\mathbb{E}[\mathbb{E}[g|C]|x]$ とすると局所 Z-推定量が得られる．条件付き期待値をとる操作は，情報幾何的射影とみなせる．射影をとる操作と推定量の漸近分散の関係について考察し，局所化による情報量損失を定量化する．また，ランダム近傍系から定義される局所 Z-推定量の漸近分散の下限を導出する．さらに，擬似尤度や合成尤度に対する既存の結果との関連について考察する．とくに，近傍の包含関係と漸近分散との関連 [2] や，Fisher 有効性をもつ局所 Z-推定量 [3] について，上記の枠組で得られた結果を紹介する．

参考文献

[1] T. Kanamori, Efficiency Bound of Local Z-Estimators on Discrete Sample Spaces, *Entropy*, 18, 273; doi:10.3390/e18070273, 2016.

[2] Liang, P. & Jordan, M.I., An Asymptotic Analysis of Generative, Discriminative, and Pseudo-likelihood Estimators. In Proceedings of the 25th International Conference on Machine Learning, pp. 584–591. 2008.

[3] Mardia, K.V., Kent, J.T., Hughes, G., & Taylor, C.C., Maximum Likelihood Estimation using Composite Likelihoods for Closed Exponential Families. *Biometrika*, 96, 975–982, 2009.