カーネルモード推定における最適カーネルの設計

京都大学大学院情報学研究科　　山崎 遼也

京都大学大学院情報学研究科　　田中 利幸

確率密度関数 f を持つ確率分布に従う i.i.d. サンプル $\{X_i \in \mathbb{R}\}_{i=1}^n$ が与えられたとする．カーネルモード推定 (Parzen, 1962) では，カーネル密度推定による密度推定結果 f_n が最大値をとる点 θ_n をモード θ の推定量 (カーネルモード推定量, KME) とする：

$$\theta_n \coloneqq \arg\max_{x\in\mathbb{R}} f_n(x),\quad \text{ただし，} f_n(x) = \frac{1}{n}\sum_{i=1}^n K_{h_n}(x - X_i) = \frac{1}{nh_n}\sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right).$$

カーネル関数 K が q 次モーメント条件

$$B_i \coloneqq \int x^i K(x)\,dx = 1\ \ (i=0);\quad 0\ \ (i=1,\ldots,q);\quad \text{非ゼロの有限値}\quad (i=q)$$

を満たすとき，KME θ_n の漸近平均二乗誤差 (AMSE) は

$$\mathrm{E}[(\theta_n - \theta)^2 | h_n] = \underbrace{\frac{h_n^{2q}\{f^{(q+1)}(\theta)\}^2 B_q^2}{(q!)^2\{f^{(2)}(\theta)\}^2}}_{\text{漸近二乗バイアス由来}} + \underbrace{\frac{f(\theta)V}{nh_n^3\{f^{(2)}(\theta)\}^2}}_{\text{漸近分散由来}} + O(h_n^{2q+2} + (nh_n)^{-1})$$

となる (ただし，$V \coloneqq \int \{K^{(1)}(x)\}^2\,du < \infty$ である)．AMSE を最小化する最適な帯域幅 h_n^{opt} と，その帯域幅を用いたときの AMSE(の主要項) は

$$h_n^{\mathrm{opt}} = \left[\frac{3(q!)^2 f(\theta)V}{2qn\{f^{(q+1)}(\theta)\}^2 B_q^2}\right]^{\frac{1}{2q+3}};\quad \mathrm{E}[(\theta_n - \theta)^2 | h_n^{\mathrm{opt}}] \propto n^{-\frac{2q}{2q+3}} \cdot \left(B_q^6 \cdot V^{2q}\right)^{\frac{1}{2q+3}}$$

となる．従って，最適な帯域幅 h_n^{opt} に対する AMSE に現れるカーネル依存項 (汎関数) $B_q^6 \cdot V^{2q}$ を最小化するカーネル関数を用いることで，KME θ_n の AMSE を最大限最小化することができる．変分問題

$$\min_K \quad B_q^6 \cdot V^{2q},\quad \text{s.t.}\quad K \text{ が } q \text{ 次モーメント条件を満たす}$$

はそのままでは解を持たないが，Granovsky and Müller (1991) はカーネル関数の符号の変化回数に制約を与えることで，偶数次数 q に対して最適なカーネルを与えた．

本研究では，(Granovsky and Müller, 1991) と同様の議論から，多変量カーネルモード推定に対する最適カーネルを導出した．また，同様の変分問題が他のカーネルベースのモード統計手法 (例えば，シンプルモード推定 (Abraham et al., 2004) やモード線形回帰 (Yao and Li, 2014)) にも現れることが分かった．これらを含む最適カーネルについて得られた結果を発表する予定である．

参考文献

C. Abraham, G. Biau, and B. Cadre. On the asymptotic properties of a simple estimate of the mode. *ESAIM: Probability and Statistics*, 8:1–11, 2004.

B. L. Granovsky and H.-G. Müller. Optimizing kernel methods: a unifying variational principle. *International Statistical Review/Revue Internationale de Statistique*, 59(3):373–388, 1991.

E. Parzen. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33 (3):1065–1076, 1962.

W. Yao and L. Li. A new regression model: modal linear regression. *Scandinavian Journal of Statistics*, 41(3): 656–671, 2014.