

# Divergence に基づく局所密度推定

島根大学総合理工学研究科 川村健太

島根大学総合理工学研究科 内藤貫太

本講演では、Bregman divergence と核関数を用いた局所最適化により構成される密度推定量とその漸近的挙動について報告する。

$U$  を凸関数とし、 $u$  をその導関数とする。また、 $f$  を  $\mathbb{R}^d$  上の未知の密度関数とし、 $f$  を推定するために用いるパラメトリックモデルを指数型分布族  $\{g_\theta(x) : \theta \in \Theta \subset \mathbb{R}^p\}$  とする。 $f$  と  $g_\theta$  の違いは、Bregman divergence

$$D_U(f, g_\theta) = \int_{\mathbb{R}^d} \{U(f(x)) - U(g_\theta(x)) - u(g_\theta(x))(f(x) - g_\theta(x))\} dx$$

で測られる。パラメトリックなアプローチでは、

$$\theta_0 = \arg \min_{\theta \in \Theta} D_U(f, g_\theta)$$

の推定が問題となる。点  $t \in \mathbb{R}^d$  における  $f(t)$  を推定するために、 $t$  に依存させる形で  $\theta_0$  の推定量を構成する。このことにより、 $g_\theta(\cdot)$  の局所的適応能力の増大が期待される。このような局所化は、 $D_U(f, g_\theta)$  に核関数を組み込むことにより実装される。考える局所化した Bregman divergence は

$$D_{U,t}(f, g_\theta) = \int_{\mathbb{R}^d} K\left(\frac{x-t}{h}\right) \{U(f(x)) - U(g_\theta(x)) - u(g_\theta(x))(f(x) - g_\theta(x))\} dx$$

となる。ここで、 $K(z)$  は原点对称かつ単峰な可積分関数であり、 $h$  は Bandwidth として局所化の程度を制御する。特に  $K(0) = 1$  のとき、 $h \rightarrow \infty$  において、 $D_{U,t}(f, g_\theta)$  は  $D_U(f, g_\theta)$  に漸近することに注意する。実際には、 $f$  からのデータ  $X_1, \dots, X_n$  が与えられたとき、 $t \in \mathbb{R}^d$  における  $f(t)$  の局所密度推定量は

$$\hat{f}(t) = \frac{g_{\hat{\theta}(t)}(t)}{\int_{\mathbb{R}^d} g_{\hat{\theta}(s)}(s) ds}$$

で与えられる。ここで、

$$\hat{\theta}(t) = \arg \min_{\theta \in \Theta} \ell_{t,n}(\theta)$$

であり、

$$\ell_{t,n}(\theta) = -\frac{1}{n} \sum_{i=1}^n K\left(\frac{X_i - t}{h}\right) u(g_\theta(X_i)) - \int_{\mathbb{R}^d} K\left(\frac{x-t}{h}\right) \{U(g_\theta(x)) - u(g_\theta(x))g_\theta(x)\} dx$$

は、 $D_{U,t}(f, g_\theta)$  の経験版である。

本講演では特に、 $\hat{f}(t)$  の挙動を支配する  $\hat{\theta}(t)$  の挙動について、 $n \rightarrow \infty$ ,  $h \rightarrow \infty$  という漸近的枠組みにおいて理論的考察を与える。