P10: Uロス関数に基づく密度推定のための ブースティング

小森 理 統計数理研究所 内藤 貫太島根大学

江口 真透 統計数理研究所

ブースティングアルゴリズム

(1) 初期値 fo を

$$\hat{L}_{\beta}(f_0) \leq \inf_{g \in \mathcal{D}_{\beta}} \hat{L}_{\beta}(g) + \epsilon$$

のように取る. 但し $\epsilon > 0$ は近似限界とする.

(2) k=1,...,K に対し、 f_k を以下のように取る.

$$f_k = \left\{ (1 - \pi_k) f_{k-1}^{1+\beta} + \pi_k \psi_k \right\}^{1/(1+\beta)}.$$

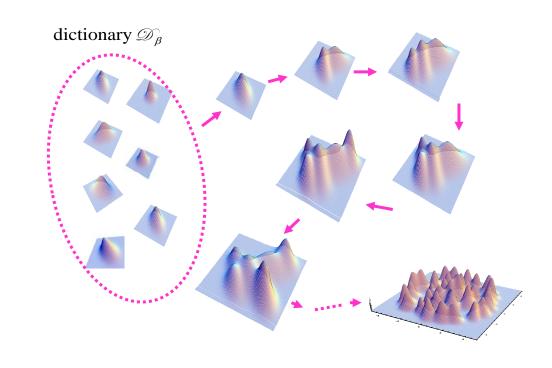
但し、 ψ_k は

 $\hat{L}_{\beta}(\psi_{k})$

$$\leq \inf_{\psi \in \mathcal{D}_{\beta}} \hat{L}_{\beta} \left(\left\{ (1 - \pi_k) f_{k-1}^{1+\beta} + \pi_k \psi \right\}^{1/(1+\beta)} \right) + \pi_k \epsilon.$$

のように取る.

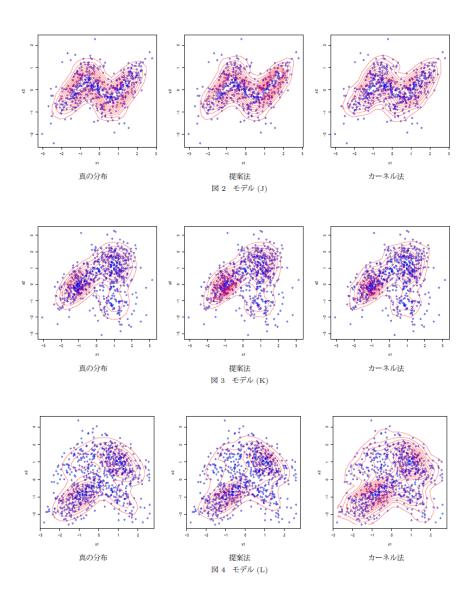
(3) 最後に $\hat{f} = f_K$ とする.



提案法とDuogn and Hazelton (2003) のカーネル法との比較を, Wand and Jones (1993)のモデル(J), (K), (L)で行った(右図). また非漸近的誤差限界が, 関数の近似性能, 推定誤差, 反復効果の3つの項で抑えられることを示した.

$$\begin{split} & \mathbb{E}_{f} D_{\beta}(f, \hat{f}) \\ & \leq \operatorname{FA}(f, \mathcal{D}_{\beta}) + \operatorname{EE}(f, \mathcal{D}_{\beta}) + \operatorname{IE}(K) \end{split}$$

但し、f は真の密度、 \hat{f} はブースティングによって推定された密度 D_{β} は β ダイバージェンス. \mathcal{D}_{β} はブースティングで使う弱判別機の集合(辞書). K はブースティングの反復回数.



確率密度の等高線