統計数理(2010)
第 58 巻 第 1 号 127–130
[研究ノート]

射影推定量についての一注意

西山 陽一[†]

(受付 2010 年 3 月 31 日;改訂 4 月 15 日;採択 4 月 15 日)

要 旨

$[-\pi,\pi]$ 上の密度に対する射影推定量の L_2 リスクの漸近限界の主要定数(leading constant)が真の密度に依存せず $1/\pi$ と陽に与えられるという注意を与える.この意味において,射影推定量にはカーネル推定量にないメリットがある.

キーワード: 密度推定,L_2 リスク,正規直交系.

1. 序

f は $[-\pi,\pi]$ 上の確率密度であるとする.それが p 回微分可能であると仮定し,その導関数を $f^{(m)}, m=0,1,\ldots,p$ と記し,さらに $||f^{(p)}||<\infty$ および $f^{(m)}(-\pi)=f^{(m)}(\pi), m=0,1,\ldots,p-1$ であることを仮定する.ただし $||\cdot||$ はルベーグ測度に関する L_2 ノルムを表す.

そのような密度をもつ分布からの i.i.d. サンプル $\{X_1,\ldots,X_n\}$ が与えられたとする.我々の目的は,f に対する**射影推定量**(projection estimator) $\widehat{f}_n$ が次のような厳密限界および漸近限界をもつという注意を与えることである.

定理 1. $\widehat{f}_n$ は以下の(1.1)によって与えられるような f に対する推定量であるとする.ただし(1.1)における正規直交系 $\{e_j\}$ は三角関数系(3.1)であるとする.

(i) 全ての n に対し,

$$\mathrm{E}||\widehat{f}_n - f||^2 \le n^{-2p/(2p+1)}\left|\frac{1}{\pi} + 2^{2p}||f^{(p)}||^2\right|.$$

(ii) $n\to\infty$ とするとき,

$$\mathrm{E}||\widehat{f}_n - f||^2 \le n^{-2p/(2p+1)}\left|\frac{1}{\pi} + o(1)\right|.$$

我々の結果は Parzen (1962) によって導入されたカーネル推定量 $\widetilde{f}_n$ と比較されるべきである.カーネル関数を K とする.よく知られた結果は次のような形をしている:

(i) 全ての n に対し,

$$\mathrm{E}||\widetilde{f}_n - f||^2 \le n^{-2p/(2p+1)} C_1(f,K);$$

(ii) $n\to\infty$ とするとき,

$$\mathrm{E}||\widetilde{f}_n - f||^2 \le n^{-2p/(2p+1)} |C_2(f,K) + o(1)|.$$

[†] 統計数理研究所:〒190–8562 東京都立川市緑町 10–3

ここで $C_1(f,K)$ および $C_2(f,K)$ は $||f^{(p)}||$ および K に依存するような定数である（もちろん $C_1(f,K)\geq C_2(f,K)$ である)．収束率 $n^{-2p/(2p+1)}$ は最適であると知られている．詳しくは例えば Tsybakov (2009) や van der Vaart (1998) を見よ．我々の結果は特に漸近的問題(ii)に対して利点をもつ．カーネル密度推定量に対する主要定数(leading constant) $C_2(f,K)$ は未知の $||f^{(p)}||$ に依存しているのに対し，我々の結果におけるそれは $1/\pi$ であって f に依存しない．統計家は真の f を事前には知らないのであるから，このような明示的な主要定数 $1/\pi$ を得ることは大きなメリットをもつ．以下の議論からわかるように，もし密度 f のサポートがある定数 $M>0$ に対し $[-M\pi, M\pi]$ であるときには，主要定数は $1/M\pi$ となる．よって我々の方法は f がコンパクトな台をもつような場合を完全にカバーしている(f の台が $[-M\pi, M\pi]$ に含まれるような大きな $M\geq 1$ を選べばよい)．

上で述べた我々の理論的結果は 1 次元の場合のみに対するものであるが，推定量の定義自体は(多次元の場合も含め)一般の状態空間 $\mathcal{X}$ に対するもので与えておく．$\{X_i\}$ は測度空間 $(\mathcal{X},\mathcal{A},\mu)$ 上で定義された(μ に関する)密度 f をもつような分布からの i.i.d. サンプルであるとする．空間 $L_2(\mathcal{X},\mu)$ を考え，内積を $\langle,\rangle$ と表し，ノルムを $||\cdot||$ と表し，正規直交系 $\{e_j\}$ をとる．関係

$$f(x)=\sum_j \langle f,e_j\rangle e_j(x), \quad f_n(x):=\sum_{j\leq d_n}\langle f,e_j\rangle e_j(x),$$

に基づいて，我々は推定量

$$\widehat{f}_n(x)=\frac{1}{n}\sum_{i=1}^{n}\sum_{j\leq d_n} e_j(X_i)e_j(x) \tag{1.1}$$

を提案する．ここで d_n は有限定数であって $n\to\infty$ とするとき発散するようなものとする．主結果においては $d_n=n^{1/(2p+1)}$ ととる．以下，2 節において主要項 $\mathrm{E}||\widehat{f}_n-f_n||^2$ を分析し，3 節においては状態空間を $\mathcal{X}=[-\pi,\pi]$ と特殊化した上でバイアス項 $||f-f_n||$ を考察する．次元 d_n はカーネル密度推定量におけるバンド幅の役割を果たす．射影推定量の構成のアイデアそのものは古くから知られているが(少なくとも Cencov, 1962 に遡る．最近の教科書としては Tsybakov, 2009 を見よ)，我々の与えた注意は現在まで指摘されていなかったものである．なお射影推定量とカーネル推定量の誤差上界のオーダーは同一となるが，この事実自体およびそれが最適率であることはよく知られていたことである．

2. 主要項

主要項については，問題を一般の状態空間 $\mathcal{X}$ のまま議論することができる．

$$\begin{aligned}
\mathrm{E}||\widehat{f}_n-f_n||^2 &= \mathrm{E}\int_{\mathcal{X}}\left|\frac{1}{n}\sum_{i=1}^{n}\sum_{j\leq d_n}(e_j(X_i)-\mathrm{E}e_j(X_i))e_j(x)\right|^2\mu(dx)\\
&= \frac{1}{n}\int_{\mathcal{X}}\sum_{j\leq d_n}\mathrm{Var}(e_j(X_1))e_j(x)^2\mu(dx)\\
&= \frac{1}{n}\sum_{j\leq d_n}\mathrm{Var}(e_j(X_1))\\
&\leq \frac{d_n}{n}\sup_j \mathrm{Var}(e_j(X_1)).
\end{aligned}$$

3. バイアス項

この節では $\mathcal{X}=[-\pi,\pi]$ とする．正規直交系 $\{e_j\}$ を

$$\frac{1}{\sqrt{2\pi}},\quad \frac{1}{\sqrt{\pi}}\cos kx,\ \frac{1}{\sqrt{\pi}}\sin kx,\quad (k=1,2,\dots) \tag{3.1}$$

ととる．j と k の関係は，コサインについては $j=1+(2k_j^c-1)$ であり，サインについては $j=1+2k_j^s$ である．よって一般に $(j-1)/2\le k_j$ である．

部分積分の公式により，例えば

$$\begin{aligned}\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\cos kx dx &= -\frac{1}{\sqrt{\pi}k}\int_{-\pi}^{\pi} f^{(1)}(x)\sin kx dx\\ &= -\frac{1}{\sqrt{\pi}k^2}\int_{-\pi}^{\pi} f^{(2)}(x)\cos kx dx\\ &= \dots\end{aligned}$$

が成り立つ．よって容易に

$$\begin{aligned}||f-f_n||^2 &= \sum_{j>d_n}\left|\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\cos k_j^c x dx\right|^2+\left|\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\sin k_j^s x dx\right|^2\\ &\le \frac{1}{|d_n/2|^{2p}}\sum_{j>d_n}\left|\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f^{(p)}(x)\sin k_j x dx\right|^2+\left|\frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f^{(p)}(x)\cos \widetilde{k}_j x dx\right|^2\\ &=: d_n^{-2p}D_n(f,p)\end{aligned}$$

が得られる．ただし 2 行目の式における k_j および $\widetilde{k}_j$ はそれぞれ k_j^c および k_j^s のいずれかを表す（いずれを表すかは p が偶数か奇数かによる）．定数 $D_n(f,p)$ は $n\to\infty$ とするときゼロに収束し，また常に $2^{2p}||f^{(p)}||^2$ より小さいものであることに注意しよう．

4. 主定理の証明

主結果を得るためには $d_n/n=d_n^{-2p}$ とおけばよい．このとき $d_n=n^{1/(2p+1)}$ である．主要定数は $\sup_j \mathrm{Var}(e_j(X_1))\le 1/\pi$ によって押さえられる．

謝　辞

この研究は日本学術振興会からの科学研究費補助金・基盤研究(C)，21540157 によって支援されたものである．

参 考 文 献

Cencov, N. N. (1962). Estimation of unknown probability density based on observations, *Doklady Akademii Nauk SSSR*, **147**, 45–48.

Parzen, E. (1962). On estimation of a probability density function and mode, *Annals of Mathematical Statistics*, **33**, 520–531.

Tsybakov, A. B. (2009). *Introduction to Nonparametric Estimation*, Springer, New York.

van der Vaart, A. W. (1998). *Asymptotic Statistics*, Cambridge University Press, Cambridge.

A Remark on Projection Estimator

Yoichi Nishiyama

The Institute of Statistical Mathematics

We remark that the leading constant of the asymptotic bound for the L_2 risk of the projection estimator for a density on $[-\pi, \pi]$ is $1/\pi$, which does not depend on the true density. In this sense, the projection estimator has a merit that the kernel estimators do not have.

Key words: Density estimation, L_2 risk, orthonormal system.