Statistics and Probability Letters 131 (2017) 46–50

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Higher order kernel density estimation on the circle ☆

Yasuhito Tsuruta [a,*], Masahiko Sagae [b]

[a] Graduate School of Human and Socio-Environment Studies, Kanazawa University, Kakumamachi Kanazawa, Ishikawa, 920-1192, Japan
[b] School of Economics, Kanazawa University, Kakumamachi Kanazawa, Ishikawa, 920-1192, Japan

ARTICLE INFO

Article history:
Received 5 December 2016
Received in revised form 29 July 2017
Accepted 4 August 2017
Available online 15 August 2017

Keywords:
Directional statistics
Circular data

ABSTRACT

A new class of pth-order kernels corresponding to new moments on the circle is introduced. We propose two methods for constructing higher-order kernel density estimators, and we derive theoretical and empirical results for the kernel density estimators.

1. Introduction

Given a random sample on the circle $\Theta_1, \ldots, \Theta_n \in [-\pi, \pi)$ from an unknown density $f(\theta)$. The kernel density estimator (KDE) on the circle is defined as

$$\hat{f}_\kappa(\theta) = \frac{1}{n}\sum_{i=1}^{n} K_\kappa(\theta - \Theta_i),$$

where $K_\kappa(\theta)$ is a symmetric kernel function, and κ is a concentration parameter that plays the role of a smoothing parameter, and corresponds to the inverse of bandwidth on the real line.

Now, we describe the theoretical properties of a standard KDE on the real line. The standard KDE is given by $\hat{f}_h(x) := n^{-1}\sum_i K_h(x - X_i)\, x \in \mathbb{R}$, where h is a bandwidth and $K_h(x - X_i) := h^{-1}K((x - X_i)/h)\, x \in \mathbb{R}$ is a symmetric kernel function on the real line. The criterion for the error between the standard KDE and the density on the real line $f(x)$ is the mean integrated squared error (MISE): $\text{MISE}[\hat{f}_h] := \int_{\mathbb{R}} \text{MSE}[\hat{f}_h(x)]dx$ where, $\text{MSE}[\hat{f}_h(x)] := \text{E}_f[\{\hat{f}_h(x) - f(x)\}^2]$ is the mean squared error. The bias of the standard KDE is expressed as

$$\text{bias}_f[\hat{f}_h(x)] = \sum_{t=1}^{p/2} h^{2t}\alpha_{2t}(K)f^{(2t)}(x)/\{(2t)!\} + o(h^p).$$

We introduce a standard pth-order kernel function $K_{(p)}$ on the real line, such that $\alpha_0(K_{(p)}) = 1,\ \alpha_0(K_{(p)}) = 0\ 0 < t < p,$ and $\alpha_p(K_{(p)}) \neq 0$, where $\alpha_t(K) = \int_{\mathbb{R}} K(x)x^t dx$ is the tth moment of the kernel on the real line $K(x)$. This improves the bias that is $\text{bias}_f[\hat{f}_h(x)] = h^p\alpha_p(K)f^{(p)}(x)/(p!) + o(h^p)$. However, no pth-order kernel functions $K_{(p)}$ can improve the variance $Var_f[\hat{f}_h(x)] = O((nh)^{-1})$. When a pth-order kernel function $K_{(p)}$ is employed, these results imply that the convergence rate of the MISE for the standard KDE is $O(n^{-2p/(2p+1)})$.

☆ It is available to refer details of some proofs into supplementary materials (see Appendix A) for convenience of explanation regarding theorem and methods.

* Corresponding author.
E-mail addresses: tsuruta9764@gmail.com (Y. Tsuruta), sagae.masahiko@gmail.com (M. Sagae).

http://dx.doi.org/10.1016/j.spl.2017.08.003
0167-7152/

The authors of Di Marzio et al. (2011) defined a pth sin-order kernel function as a kernel function that satisfies $\eta_j(K_\kappa) = 0$ for $0 < j < p$ and $\eta_p(K_\kappa) \neq 0$, where $\eta_j(K_\kappa) := \int_{-\pi}^{\pi} \sin^j(\theta) K_\kappa(\theta) d\theta$ are jth sin-order moments. They derived the MISE for the KDE on the circle $\hat{f}_\kappa$: $\mathrm{MISE}[\hat{f}_\kappa(\theta)] := \int_{-\pi}^{\pi} \mathrm{MSE}[\hat{f}_\kappa(\theta)] d\theta$, where $\mathrm{MSE}[\hat{f}_\kappa(\theta)] := \mathrm{E}_f[\{\hat{f}_\kappa(\theta) - f(\theta)\}^2]$, and showed that the convergence rate of the MISE for the von Mises (VM) kernel (the second sin-order kernel) is $O(n^{-4/5})$. They argued that sin-order kernel functions do not necessarily yield smaller biases. In other words, the order p of the pth sin-order kernel function does not generally correspond to the convergence rate of the MISE. For example, Tsuruta and Sagae (2017) indicated that the rate of the wrapped Cauchy kernel (the second sin-order kernel) for the optimal MISE's rate is $O(n^{-2/3})$. They succeeded in constructing higher-order kernels and improving the bias through "Twicing": the bias reduction technique.

In the next section, we introduce a new class of pth-order kernel functions with new moments. The convergence rate of the MISE of $\hat{f}_\kappa$ using such a kernel is $O(n^{-2p/(2p+1)})$. In addition, $\hat{f}_\kappa$ also has the property of asymptotic normality.

We show that higher-order KDEs can be constructed by applying either the additive method from Jones and Foster (1993) or the multiplicative method from Terrell and Scott (1980) to our kernel functions. The two methods are standard ways of constructing higher-order KDEs on the real line, and are frequently employed in practical analyses. Our simulation demonstrates that these higher-order KDEs exhibit better properties than the second-order KDE as n becomes larger.

2. Properties of higher-order kernel density estimators

We stand with the definition of a kernel function for circular data as proposed by Hall et al. (1987).

Definition 1. A function $K_\kappa(\theta) : [-\pi, \pi) \to \mathbb{R}$ is said to be a kernel function. Let $K_\kappa(\theta)$ denote $K_\kappa(\theta) := C_\kappa^{-1}(L) L_\kappa(\theta)$, where $L_\kappa(\theta) := L(\kappa\{1 - \cos(\theta)\})$ and $C_\kappa(L) := \int_{-\pi}^{\pi} L_\kappa(\theta) d\theta$. We define the lth moment of L as

$$\mu_l(L) := \int_0^\infty L(r) r^{(l-1)/2} dr,$$

where $l \geq 0$ is even and $r = \kappa\{1 - \cos(\theta)\}$. The main term L satisfies the following four conditions.

(i) The derivative $L'(r) := dL(r)/dr$ is continuous.
(ii) $L(r) r^{(v+1)/2} \to 0$ as $r \to \infty$.
(iii) The term $\delta_{2t}(L) := \int_{-\infty}^{\infty} L^2(z^2/2) z^{2t} dz$ is bounded for $t = 0, 1$.
(iv) Let $v \geq 0$ denote any even number. Then, $\mu_l(L)$ is bounded for $0 \leq l \leq v$ and $\mu_l(L) = \mu_{\kappa,l}(L) + O(\kappa^{-(v+2)/2})$, where $\mu_{\kappa,l}(L) := \int_0^\kappa L(r) r^{(l-1)/2} dr$.

The condition (i) and (ii) are required to obtain the partial integral $\int^\infty L'(r) r^{(v+1)/2} dr = -(j+1)\mu_v(L)/2$. The term $\delta_{2t}(L)$ represents the constant of the variance of $\hat{f}_\kappa$. The condition (iv) is employed to derive the bias of $\hat{f}_\kappa$.

We now define the pth-order kernel function.

Definition 2 (*pth-order Kernel Function*). Let $p \geq 2$ be an even number. Then, we say that $K_\kappa(\theta)$ is a pth-order kernel if $v \geq p + 2$, and

$$\mu_0(L) \neq 0, \quad \mu_l(L) = 0, \quad l = 2, 4, \ldots, p-2 \quad \text{and} \quad \mu_l(L) \neq 0 \quad l = p.$$

We have the following result for determining properties for $\hat{f}_\kappa$ using a pth-order kernel function.

Theorem 1. *Assume that the following conditions hold:*

(i) $\kappa = \kappa(n)$ *and* $\lim_{n\to\infty} \kappa(n) = \infty$,
(ii) $\lim_{n\to\infty} n^{-1} \kappa^{1/2}(n) = 0$,
(iii) f *is* $(p+2)$*th differentiable and* $f^{(s)}$ *is square-integrable,* $s = 1, 2, \ldots, p$,
(iv) K_κ *is a* p *th-order kernel function.*

Then, MISE is given by

$$MISE[\hat{f}_\kappa] = \frac{\mu_p^2(L)}{\mu_0^2(L)} R\left(\sum_{t=1}^{p/2} \frac{b_{p,2t} f^{(2t)}}{2t!} \right) \kappa^{-p} + n^{-1}\kappa^{1/2} d(L) + o(\kappa^{-p} + n^{-1}\kappa^{1/2}).$$

The first two terms of the above are referred as $AMISE[\hat{f}_\kappa]$*. The minimizer* κ^* *of* $AMISE[\hat{f}_\kappa]$ *is equal to*

$$\kappa^* = \left[\frac{2p\mu_p^2(L) R(\sum_{t=1}^{p/2} [b_{p,2t} f^{(2t)}/(2t!)]) n}{\mu_0^2(L) d(L)} \right]^{2/(2p+1)}. \tag{1}$$

Thus, the optimal MISE for the KDE $\hat{f}_\kappa$ *with a* p *th-order kernel is* $O(n^{-2p/(2p+1)})$*.*