Proceedings Paper:
Di Marzio, M, Fensore, S, Panzera, A et al. (2020) Kernel Circular Deconvolution Density Estimation. In: ISNPS 2018. 4th ISNPS: Conference of the International Society for Non-Parametric Statistics, 11-15 Jun 2018, Salerno, Italy. Springer International Publishing, pp. 183-191. ISBN: 9783030573058. ISSN: 2194-1009. EISSN: 2194-1017.

https://doi.org/10.1007/978-3-030-57306-5_17

Kernel circular deconvolution density estimation

Marco Di Marzio, Stefania Fensore, Agnese Panzera, Charles C. Taylor

Abstract We consider the problem of nonparametrically estimating a circular density from data contaminated by angular measurement errors. Specifically, we obtain a kernel-type estimator with weight functions that are reminiscent of deconvolution kernels. Here, differently from the Euclidean setting, discrete Fourier coefficients are involved rather than characteristic functions. We provide some simulation results along with a real data application.

1 Introduction

Circular data arise when the sample space is described by a unit circle. If compared to a linear scale, the main features of circular observations are that the beginning and the end of the measurement scale coincide, and their common location is called the origin (or zero direction) which is arbitrarily chosen. Once the origin and the sense of rotation have been chosen, any circular observation can be measured by an angle ranging, in radians, from 0 to 2π. Circular data often arise in biology, meteorology and geology; other examples include phenomena that are periodic in time. For comprehensive accounts of statistics for circular data see, for example, [4] and [5].

Marco Di Marzio
University of Chieti-Pescara, e-mail: mdimarzio@unich.it

Stefania Fensore
University of Chieti-Pescara, e-mail: stefania.fensore@unich.it

Agnese Panzera
University of Florence, e-mail: agnese.panzera@unifi.it

Charles C. Taylor
University of Leeds, e-mail: charles@maths.leeds.ac.uk

In this paper we discuss the problem of nonparametrically estimating a circular density when data are observed with error. Specifically, here we consider the case of measurement errors described by circular random variables. This problem has been studied by [3] who proposed an estimator constructed as a truncated development of the density to be estimated represented by trigonometric basis where the theoretical coefficients are replaced by empirical ones. Then [1], using a model selection procedure, derived an adaptive penalized contrast estimator, and [6] proposed an orthogonal series estimator optimal in the minimax sense.

In the Euclidean setting the problem of estimating a density in the context of errors-in-variables has been widely pursued. The most popular method is a nonparametric one based on kernel-type estimators. For an exhaustive treatment of density estimation with errors-in-variables and related topics see [2] and the references therein. In the directional setting the kernel-based methods for errors-in-variables problems seem to be substantially unexplored. In this article we propose to extend this approach to the estimation of a circular density.

After recalling in Section 2 some preliminaries about Fourier series and nonparametric estimation of circular densities in the error-free case, in Section 3 we discuss the extension of the kernel-type density estimator to the case where variables are observed with errors. Then, in Section 4 we present some simulation results, and in Section 5 we end up with a real data application.

2 Preliminaries

In this section we provide some basic facts about Fourier series representation of circular densities and recall the definition of the circular kernel density estimator.

2.1 Trigonometric moments and Fourier series

Let Q be a circular random variable and denote by f_Q its probability density function. Due to the periodic nature of Q, its distribution is the same as the distribution of $Q+2\pi$; this implies that the characteristic function of Q, which is

$$\varphi_Q(\ell) = E[e^{i\ell Q}] = \int_0^{2\pi} e^{i\ell q} f_Q(q)dq,$$

is defined only at integer ℓs. Moreover, for any $\ell \in \mathbb{Z}$, one has

$$|\varphi_Q(\ell)| \leq 1, \quad \varphi_Q(0) = 1, \quad \bar{\varphi}_Q(\ell) = \varphi_Q(-\ell),$$

where $\bar{\varphi}_Q(\cdot)$ stands for the complex conjugate of $\varphi_Q(\cdot)$. Notice that the complex numbers $\{\varphi_Q(\ell), \ell \in \mathbb{Z}\}$ are the coefficients in the Fourier series representation (in complex form) of f_Q and correspond to the *trigonometric moments* of Q about the mean direction, i.e., letting

$$\alpha_\ell = E[\cos(\ell Q)],\ \ \beta_\ell = E[\sin(\ell Q)],$$

it holds that $\varphi_Q(\ell) = \alpha_\ell + i\beta_\ell$; clearly, for any $\ell \in \mathbb{Z}$,

$$\alpha_{-\ell} = \alpha_\ell, \quad \beta_{-\ell} = -\beta_\ell, \quad |\alpha_\ell| \le 1, \quad |\beta_\ell| \le 1.$$

Then, assuming that f_Q is a square integrable function on $[0, 2\pi)$, for $q \in [0, 2\pi)$, one can recover $f_Q(q)$ from the Fourier coefficients by the expansion

$$f_Q(q) = \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \varphi_Q(\ell) \exp(-i\ell q) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{\ell=1}^{\infty} (\alpha_\ell \cos(\ell q) + \beta_\ell \sin(\ell q)) \right\}.$$

Above equation is the analogous of the inversion formula for characteristic functions of real-valued random variable.

In the Euclidean setting, the smoothness of a density can be determined by the rate of decay of the Fourier transforms: a polynomial decay characterizes *ordinary smooth* functions, while an exponential decay characterizes *supersmooth* ones. Analogously, for a circular density the smoothness can be defined according to the rate of decay of the coefficients in its Fourier series representation.

Formally, following [3] the density of a circular random variable is supersmooth if, for $\ell \in \mathbb{Z}$,

$$c_0(|\ell| + 1)^{a_0} e^{-b|\ell|^a} \le |\varphi_Q(\ell)| \le c_1(|\ell| + 1)^{a_1} e^{-b|\ell|^a},$$

and it is ordinary smooth if

$$c_0(|\ell| + 1)^{-a_0} \le |\varphi_Q(\ell)| \le c_1(|\ell| + 1)^{-a_1},$$

where a, b, c_0, c_1 are positive constants and a_0, a_1 are real ones.

We recall that for a wrapped circular distribution, the trigonometric moment of order $\ell \in \mathbb{Z}$ corresponds to the value of the characteristic function of the unwrapped random variable, say φ_X, at (integer) ℓ, i.e. $\varphi_Q(\ell) = \varphi_X(\ell)$.

Examples of supersmooth densities include the densities of wrapped Normal and wrapped Cauchy distribution; conversely, the wrapped Laplace and the wrapped Gamma densities are examples of ordinary smooth ones. See Figure 1 for some examples of density of wrapped distributions.

2.2 *Circular density estimation in the error-free case*

Given a random sample of angles $\Theta_1, \ldots, \Theta_n$ from an unknown circular density f_Θ, the kernel estimator of f_Θ at $\theta \in [0, 2\pi)$ is defined as

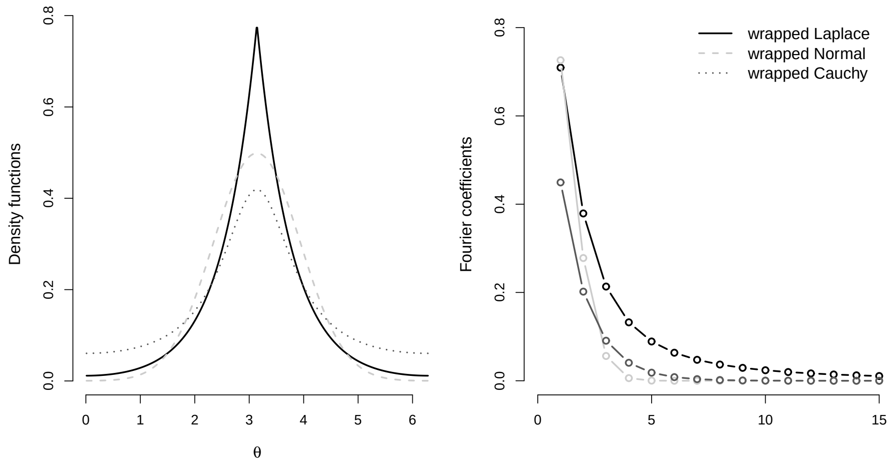

Fig. 1 Examples of wrapped densities sharing the values of mean and variance of their un-wrapped versions (left) and corresponding Fourier coefficients (right).

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

where K_κ is a circular kernel, i.e. a periodic, unimodal, symmetric density function with concentration parameter $\kappa > 0$, which admits a convergent Fourier series representation as follows

$$K_\kappa(\theta) = \frac{1 + 2\sum_{\ell=1}^{\infty} \gamma_\ell(\kappa)\cos(\ell\theta)}{2\pi}.$$

Notice that, due to the symmetry, the Fourier coefficients of K_κ satisfy $\beta_\ell = 0$ and $\alpha_\ell = \gamma_\ell(\kappa)$ for any ℓ.

Classical examples of circular kernels are the von Mises density with $\gamma_\ell(\kappa) = \mathscr{I}_\ell(\kappa)/\mathscr{I}_0(\kappa)$, where $\mathscr{I}_\ell(\kappa)$ is the modified Bessel function of order ℓ; the Wrapped Normal and Wrapped Cauchy densities where $\gamma_\ell(\kappa) = \kappa^{\ell^2}$ and $\gamma_\ell(\kappa) = \kappa^\ell$, respectively.

As it happens in the linear setting, the role of the kernel function is to emphasize, in the estimation process, the contribution of the observations which are in a neighbourhood of the estimation point. Here, the concentration parameter κ controls the width of that neighbourhood in such a way that its role is the inverse of the bandwidth in the linear case, in the sense that smaller values of κ give wider neighbourhoods.

3 Kernel density estimator in the errors-in-variables case

We consider the problem of estimating a density of a circular random variable Θ which is observed with error. In particular, we deal with the measurement error case where we wish to estimate the density f_Θ of Θ but we observe independent copies of the circular random variable

$$\Phi = (\Theta + \varepsilon)\mathsf{mod}(2\pi),$$

where ε is a random angle independent of Θ, whose density f_ε is assumed to be a known circular density symmetric around zero.

Notice that the density f_Φ of Φ is the *circular convolution* of f_Θ and f_ε, i.e., for $\theta \in [0, 2\pi)$,

$$f_\Phi(\theta) = \int_0^{2\pi} f_\Theta(\omega) f_\varepsilon(\theta - \omega) d\omega, \tag{1}$$

so, the estimation of f_Θ reduces to a circular *deconvolution* density problem.

Similarly to the Euclidean case, equation (1) implies that, for $\ell \in \mathbb{Z}$,

$$\varphi_\Phi(\ell) = \varphi_\Theta(\ell)\varphi_\varepsilon(\ell),$$

then, a naive estimator of f_Θ at $\theta \in [0, 2\pi)$ could be

$$\tilde{f}_\Theta(\theta) = \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \frac{\hat{\varphi}_\Phi(\ell)}{\varphi_\varepsilon(\ell)} e^{-i\ell\theta}, \tag{2}$$

where $\hat{\varphi}_\Phi(\ell) = \frac{1}{n}\sum_{j=1}^{n} e^{i\ell\Phi_j}$ is the empirical version of $\varphi_\Phi(\ell)$. Now, since

$$\int_{-\pi}^{\pi} \left(f_\Theta(\theta) - \tilde{f}_\Theta(\theta)\right) d\theta = \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \left(\varphi_\Theta(-\ell) - \frac{\hat{\varphi}_\Phi(-\ell)}{\varphi_\varepsilon(-\ell)}\right) \left(\varphi_\Theta(\ell) - \frac{\hat{\varphi}_\Phi(\ell)}{\varphi_\varepsilon(\ell)}\right)$$

we have that rapid decays of $\varphi_\varepsilon(\ell)$ lead to big discrepancies between $f_\Theta(\theta)$ and $\tilde{f}_\Theta(\theta)$ even in correspondence of small discrepancies between $\varphi_\Theta(\ell)$ and $\hat{\varphi}_\Phi(\ell)$. Therefore, in order to regularize estimator (2), a possible remedy is to introduce the characteristic function of a circular kernel K_κ, say $\varphi_{K_\kappa}(\ell)$, as a damping factor, i.e.

$$\begin{aligned}
\tilde{f}_\Theta(\theta;\kappa) &= \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \frac{\hat{\varphi}_\Phi(\ell)}{\varphi_\varepsilon(\ell)} \varphi_{K_\kappa}(\ell) e^{-i\ell\theta} \\
&= \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \frac{\varphi_{K_\kappa}(\ell)}{\varphi_\varepsilon(\ell)} \frac{1}{n} \sum_{j=1}^{n} e^{i\ell\Phi_j} e^{-i\ell\theta} \\
&= \frac{1}{2\pi} \frac{1}{n} \sum_{\ell=-\infty}^{\infty} \sum_{j=1}^{n} \frac{\varphi_{K_\kappa}(\ell)}{\varphi_\varepsilon(\ell)} e^{-i\ell(\theta - \Phi_j)},
\end{aligned}$$

which leads to the following circular deconvolution estimator of $f_\Theta(\theta)$ at $\theta \in [0, 2\pi)$

$$\tilde{f}_\Theta(\theta;\kappa) = \frac{1}{2\pi}\frac{1}{n}\sum_{j=1}^{n}\left(1+2\sum_{\ell=1}^{\infty}\frac{\gamma_\ell(\kappa)}{\lambda_\ell(\kappa_\varepsilon)}\cos(\ell(\theta-\Phi_j))\right), \tag{3}$$

where $\gamma_\ell(\kappa)$ and $\lambda_\ell(\kappa_\varepsilon)$, respectively are, the ℓth coefficients in the Fourier series representation of K_κ and f_ε. Also, in order to guarantee that estimator (3) is well defined, we assume that a) the error density is an infinitely divisible distribution, i.e. it has nonvanishing Fourier coefficients $|\lambda_\ell(\kappa_\varepsilon)| > 0$ for any integer ℓ, and b) the kernel K_κ and $\tilde{f}_\Theta(\cdot;\kappa)$ are square integrable functions, i.e, using the Parseval's identity,

$$\frac{1}{2\pi}\left(1+2\sum_{\ell=1}^{\infty}\gamma_\ell^2(\kappa)\right) < \infty \quad \text{and} \quad \frac{1}{2\pi}\left(1+2\sum_{\ell=1}^{\infty}\frac{\gamma_\ell^2(\kappa)}{\lambda_\ell^2(\kappa_\varepsilon)}\right) < \infty.$$

4 Simulations

In this section we compare the performances of the deconvolution estimator and the standard kernel density one in a simulation setting.

In particular, we consider the von Mises density (vM) with mean direction and concentration respectively equal to π and 2 as the target density f_Θ, and the wrapped Laplace (wL), wrapped Normal (wN) or wrapped Cauchy (wC) with zero mean direction and different values of the concentration parameter as the error density f_ε. Notice that the concentration parameter takes non-negative real values for both vM and wL but with opposite meaning in the sense that for latter lower values of the concentration parameter give higher concentration. Differently, for both wN and wC the concentration parameter ranges from 0 to 1 with the concentration increasing with the value of the parameter.

The noise-to-signal ratio (NSR), which is defined as the ratio between the circular variance of ε and that one of Θ, ranges from 16% to 47%. The three considered settings are showed in Figure 2, where for ease of presentation the target density has been represented with zero mean direction.

We generate 500 samples of size $n = 100, 200$ and 400. We compare the estimators in terms of averaged integrated squared error (AISE). In particular we calculate the ratio $AISE_{dec}/AISE_{kde}$, where dec stands for $\tilde{f}_\Theta(\theta;\kappa)$ and kde stands for the standard circular kernel estimator $\hat{f}_\Theta(\theta;\kappa)$. The smoothing parameter κ has been selected by using the least squares cross-validation. Results are summarized in Table 1 and Figure 3. It can be seen that the deconvolution estimator outperforms the standard one especially when the NSR is moderate or the error density is ordinary smooth.

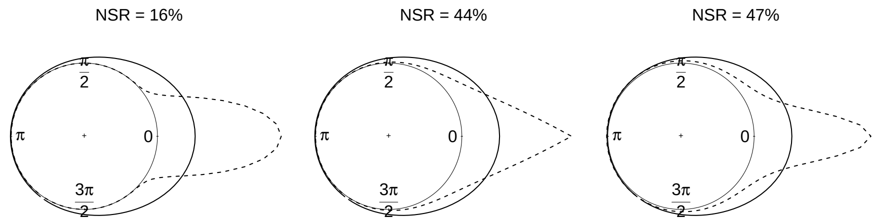

Fig. 2 von Mises density with zero mean direction and concentration parameter equals 2 (continuous) and error densities (dashed) which are wrapped Normal (left), wrapped Laplace (middle) and wrapped Cauchy (right) with zero mean direction and concentrations respectively equal to 0.97, 0.33, 0.80.

NSR	Target density	Error density	n=100	n=200	n=400
16%	vM(π, 2)	wN(0, 0.97)	0.755	0.782	0.769
44%	vM(π, 2)	wL(0, 0.33)	0.866	0.857	0.839
47%	vM(π, 2)	wC(0, 0.80)	0.966	1.015	1.085

Table 1 Comparison between the deconvolution estimator and the circular kernel density one ($AISE_{dec}/AISE_{kde}$) over 500 samples of sizes 100, 200 and 400 drawn from the target population contaminated by noise obtained by different error populations.

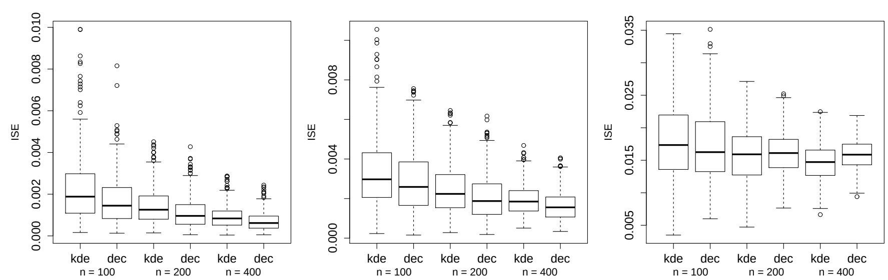

Fig. 3 Comparison between the deconvolution estimator and the circular kernel density one in terms of integrated squared errors (ISE) over 500 contaminated samples of sizes 100, 200 and 400 with a NSR equals to 16% (left), 44% (middle) and 47% (right).

5 Real data example

We consider the classical dataset described by [4] concerning the directions chosen by 100 ants in response to an evenly illuminated black target placed at π. [4] showed that classical parametric models, like von Mises, are not suited for these data. However, he concluded for an unimodal population. A nonparametric approach has been suggested by [3], who, in the context of errors-in-variables modelling, concluded for a certain evidence about multimodality. His approach is based on orthogonal trigonometric series. The rationale behind the errors-in-variables hypothesis is that, due to the typical jerky movement of the insect, the point where the ant intersects the circle can be treated as indirect observation of the direction chosen by the ant.

We compare the standard circular kernel density estimator with our deconvolution one. Specifically, we have assumed a wrapped Laplace error with zero mean and concentration equal to 0.2, employing a wrapped Normal weight function whose smoothing parameter has been selected by least squares cross-validation. As it can be seen in Figure 4 the proposed deconvolution estimator reveals the presence of three modes more efficiently than the standard method.

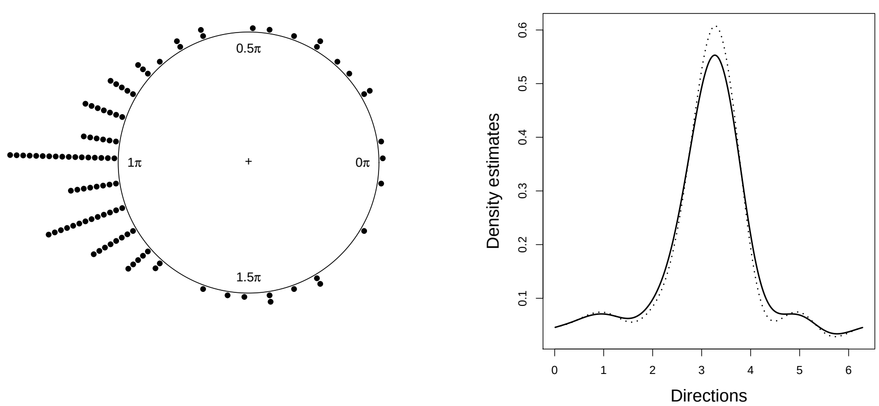

Fig. 4 Ants data (left) and kernel density estimate (continuous) and deconvolution one (dotted) of the directions of ants (right).

References

1. Comte, F., Taupin, M.L.: Adaptive density deconvolution for circular data. Prépublication MAP5 2003-10 report, Universit Paris Descarte (2003)
2. Delaigle A.: Nonparametric kernel methods with errors-in-variables: constructing estimators, computing them, and avoiding common mistakes. Aust. N. Z. J. Stat. **56**, 105–124 (2014)
3. Efromovich, S.: Density Estimation for the Case of Supersmooth Measurement Error. J. Amer. Statist. Assoc. **92**, 526–535 (1997)
4. Fisher, N.I.: Statistical analysis of circular data. Cambridge University Press (1993)
5. Jammalamadaka S.R., SenGupta A.: Topics in Circular Statistics. World Scientific (2001)
6. Jhoannes, J., Schwarz, M.: Adaptive circular deconvolution by model selection under unknown error distribution. *Bernoulli*, **19**, 1576–1611 (2013)