Kernel circular density estimation with errors in variables

Panzera Agnese [1], Di Marzio Marco [2], Fensore Stefania[3], Taylor Charles C.[4]

[1] *Dipartimento di Statistica, Informatica, Applicazioni, Università di Firenze, agnese.panzera@unifi.it*
[2] *Dipartimento di Scienze filosofiche, pedagogiche ed economico-quantitative, Università di Chieti-Pescara, marco.dimarzio@unich.it*
[3] *Dipartimento di Scienze filosofiche, pedagogiche ed economico-quantitative, Università di Chieti-Pescara, stefania.fensore@unich.it*
[4] *Department of Statistics, University of Leeds, charles@maths.leeds.ac.uk*

Résumé. On considère le problème de l'estimation non paramétrique d'une densité circulaire à partir de données contaminées par des erreurs angulaires. On propose pour la tâche un estimateur à noyau dont les poids rappellent les noyaux de déconvolution. Une étude de simulation a été réalisée pour démontrer la performance de l'estimateur proposé.
Mots-clés. Convolution circulaire, noyaux circulaires, coefficients de Fourier.

Abstract. The problem of nonparametrically estimating a circular density from data contaminated by angular errors is considered. A kernel-type estimator whose weights are reminiscent of deconvolution kernels is proposed for the task. A simulation study has been carried out to show the performance of the proposed estimator.
Keywords. Circular convolution, circular kernels, Fourier coefficients.

1 Introduction

Circular data arise when the sample space is the unit circle. In particular, a circular observation can be represented as a point on the circumference of the unit circle and, once a zero direction and a sense of rotation have been chosen, can be measured by an angle ranging, in radians, from 0 to 2π. Circular data are common in biology, meteorology and geology. For a comprehensive account of statistics for circular data see, for example, Mardia and Jupp (2009).

The problem of estimating a circular density with data corrupted by measurement errors has been studied by Efromovich (1997) who proposed an estimator constructed as a truncated trigonometric series of the target density where the theoretical coefficients are replaced by empirical ones. Comte and Taupin (2003) derived an adaptive penalized contrast estimator, while Johannes and Schwartz (2009) proposed an orthogonal series estimator optimal in the minimax sense.

In the *linear* setting the problem of estimating a density when variables are observed with errors has been widely studied. A very popular method to deal with this problem is based on kernel estimation. For an exhaustive treatment of kernel density estimation in the errors-in-variables context and related problems, see Delaigle (2014) and the references therein. In the directional setting the kernel-based methods for errors-in-variables problems seem to be substantially unexplored, and here we propose to extend this approach to circular density estimation. In Subsection 2.1 we recall some preliminaries, and then in Subsection 2.2 we briefly discuss the construction of the kernel estimator to tackle this problem. Finally, Subsection 2.3 collects some simulation results.

2 A kernel circular deconvolution estimator

2.1 Preliminaries

The characteristic function of a whatever circular random variable Θ, with absolutely continuous density f_Θ, is the sequence of complex numbers $\{\varphi_\Theta(\ell), \ell = 0, \pm 1, \pm 2, \ldots\}$, where

$$\varphi_\Theta(\ell) = \int_0^{2\pi} e^{i\ell\theta} f_\Theta(\theta) d\theta.$$

The complex number $\varphi_\Theta(\ell)$ is also referred to as the ℓth *trigonometric moment* of Θ about the zero direction, and can be expressed as $\varphi_\Theta(\ell) = \alpha_\ell + i\beta_\ell$, where

$$\alpha_\ell = \mathbb{E}[\cos(\Theta)] \quad \text{and} \quad \beta_\ell = \mathbb{E}[\sin(\Theta)].$$

Then, assuming that f_Θ is a square integrable function on $[0, 2\pi)$, for $\theta \in [0, 2\pi)$, one can recover $f_\Theta(\theta)$ from the Fourier series expansion

$$f_\Theta(\theta) = \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \varphi_\Theta(\ell) \exp(-i\ell\theta) = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{\ell=1}^{\infty} (\alpha_\ell \cos(\ell\theta) + \beta_\ell \sin(\ell\theta)) \right\}.$$

Now assume that f_Θ is unknown, and let $\Theta_1, \ldots, \Theta_n$ be a random sample of angles from f_Θ. The kernel estimator of f_Θ at $\theta \in [0, 2\pi)$ is defined as

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

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

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

Details for kernel estimation of circular density are provided by Di Marzio et al.(2011).

2.2 Deconvolution circular kernels

Now, consider an errors-in-variables density estimation problem, where we wish to estimate the circular density f_Θ of Θ but we observe n 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. We also assume that f_Θ, f_ε and f_Φ are square integrable densities on $[0, 2\pi)$ such that they admit convergent Fourier series representations.

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

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

so, the estimation of f_Θ reduces to a circular *deconvolution* density problem. The identity above implies that, for $\ell \in \mathbb{Z}$, $\varphi_\Phi(\ell) = \varphi_\Theta(\ell)\varphi_\varepsilon(\ell)$, so assuming $\varphi_\varepsilon(\ell) \neq 0$, a naive estimator of f_Θ at $\theta \in [0, 2\pi)$ could be

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

where $\hat{\varphi}_\Phi(\ell) = \frac{1}{n}\sum_{j=1}^{n} e^{i\ell\Phi_j}$ is the empirical version of $\varphi_\Phi(\ell)$. A *regularized* version of the above estimator can be constructed by using the characteristic function of a circular kernel K_κ, say $\varphi_{K_\kappa}(\ell)$, as a *damping factor*. This yields

$$\tilde{f}_\Theta(\theta; \kappa) = \frac{1}{n} \sum_{j=1}^{n} \widetilde{K}_\kappa(\Phi_j - \theta),$$

where

$$\begin{aligned}\widetilde{K}_\kappa(\Phi_j - \theta) &= \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \frac{\varphi_{K_\kappa}(\ell)}{\varphi_\varepsilon(\ell)} e^{i\ell(\Phi_j - \theta)} \\ &= \frac{1}{2\pi} \left\{ 1 + 2 \sum_{\ell=1}^{\infty} \frac{\gamma_\ell(\kappa)}{\lambda_\ell(\kappa_\varepsilon)} \cos(\ell(\Phi_j - \theta)) \right\},\end{aligned}$$

with $\gamma_\ell(\kappa)$ and $\lambda_\ell(\kappa_\varepsilon)$ respectively being the ℓth coefficients of the cosine terms in the Fourier series representation of K_κ and f_ε.

In the linear setting, the smoothness of a density can be determined by the rate of decay of the Fourier transform: a polynomial decay characterizes *ordinary smooth* functions, while an exponential decay characterizes *supersmooth* ones. Similarly, the smoothness of

NSR	Target density	Error density	n=100	n=200	n=400
5%	vM(0, 2)	wL(0, 0.1)	0.960	1.006	0.998
16%	vM(0, 8)	wL(0, 0.1)	0.995	0.958	0.915
44%	vM(0, 2)	wL(0, 0.33)	0.866	0.857	0.839
47%	vM(0, 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 target populations contaminated by noise obtained by different error populations.

a circular density can be defined according to the rate of decay of the coefficients in its Fourier series representation. Recalling that for the density of a wrapped distribution, the trigonometric moment of order ℓ corresponds to the characteristic function of the unwrapped one at (integer) ℓ, we have that examples of supersmooth densities include the densities of wrapped Normal and wrapped Cauchy distributions, while the densities of wrapped Laplace and wrapped Gamma distributions are examples of ordinary smooth circular densities. As we will see in the simulation results, the smoothness of the error density may affect the performance of the proposed estimator.

2.3 Simulations

We compare the performances of our estimator and the standard kernel density one in a simulation setting. In particular, we consider the von Mises (vM) density with zero mean direction and different values of the concentration parameter as the target density f_Θ, and the wrapped Laplace (wL) or the wrapped Cauchy (wC) with zero mean direction and different values of the concentration parameters 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 the latter one lower values of the concentration parameter give higher concentration. Differently, for 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 variances* of ε and Θ, ranges from 5% to 47%.

We generate 500 samples of sizes $n = 100, 200$ and 400. To evaluate the performances of the estimator, we calculate the averaged integrated squared error (AISE), and the ratio $AISE_{dec}/AISE_{kde}$, where *dec* and *kde* respectively stand for $\tilde{f}_\Theta(\theta; \kappa)$ and $\hat{f}_\Theta(\theta; \kappa)$. The parameter κ is selected by least squares cross-validation. As it can be seen in Table 1, the deconvolution estimator outperforms the standard one especially when the NSR is moderate or the error density is ordinary smooth.

Bibliography

Comte, F., Taupin, M.L. (2003). Adaptive density deconvolution for circular data. Prépublication MAP5 2003-10 report, Université Paris Descarte

Delaigle A. (2014). Nonparametric kernel methods with errors-in-variables: constructing estimators, computing them, and avoiding common mistakes, *Australian & New Zeland Journal of Statistics*, 56, pp. 105-124

Di Marzio M., Panzera A., Taylor, C. C. (2011). Kernel density estimation on the torus. *Journal of Statistical Planning and Inference*, 141, pp. 2156-2173

Efromovich, S. (1997). Density Estimation for the Case of Supersmooth Measurement Error, *Journal of the American Statistical Association*, 92, pp. 526-535

Johannes, J., Schwarz, M. (2013). Adaptive circular deconvolution by model selection under unknown error distribution. *Bernoulli*, 19, pp. 1576-1611

Mardia, K.V., and Jupp P.E. (2009). Directional statistics. John Wiley & Sons, 2009.