Directional–linear kernel density estimator

Eduardo García–Portugués, Rosa M. Crujeiras, and Wenceslao González–Manteiga

Departamento de Estatística e Investigación Operativa. Universidade de Santiago de Compostela.

Keywords: Directional–linear; Circular–linear; Directional data; Circular data; Kernel Density estimation.

Directional statistics deal with data lying on the q–dimensional sphere $\Omega_q = \left\{\mathbf{x} \in \mathbb{R}^{q+1} : ||\mathbf{x}|| = 1\right\}$, with the most common cases corresponding to the circle ($q = 1$, circular data) and the sphere ($q = 2$). Directional data examples appear in many applied fields, such as wind incidence direction in meteorological research (or more recently, in energy supply studies), angles in molecular structures or stars position in the sky. Statistical methods for analyzing directional data (see Mardia and Jupp, 2000) must account for their special nature, both from parametric and nonparametric approaches.

Nonparametric kernel density estimator (see Wand and Jones, 1995) provides a powerful descriptive and inferential tool. The classical kernel density estimator for linear data was adapted to the directional framework by Hall *et al.* (1987) and Bai *et al.* (1988), who proved large sample properties in terms of bias, variance and squared–error and Kullback–Leibler losses. Klemelä (2000) extended these results for the estimation of the density derivatives and Zhao and Wu (2001) derive a central limit theorem for the integrated square error of the directional kernel density estimator. Analogue to the linear case, a smoothing parameter must be selected, and for that purpose Hall *et al.* (1987) suggested data–driven cross–validation procedures.

However, in the examples mentioned above, it may be interesting to assess the relation between directional and linear random variables. For instance, in environmental protection, pollutants concentration and wind direction joint structure may be used to detect emission sources; in energy supply studies, distribution of wind direction and wind speed may be useful to determine wind mills locations in a wind farm, or in astronomy, luminosity and stars position density will provide a better understanding of the stellar objects.

In this work, an estimator for the joint density of a directional–linear random variable $(\mathbf{X}, Z)$ with support in $\Omega_q \times \mathbb{R}$ is proposed. This estimator is based on a directional–linear kernel product and expressions for bias, variance and mean square error are derived. Optimal smoothing parameters in terms of the asymptotic mean integrated square error criterion is also provided. The finite sample properties of the estimator are explored throughout a simulation study for the circular–linear and spherical–linear cases. Finally, the estimator is illustrated with real circular–linear and spherical–linear data examples.

References

[1] Z. D. Bai, C. R. Rao, and L. C. Zhao. Kernel estimators of density function of directional data. *J. Multivariate Anal.*, 27(1):24–39, 1988.

[2] P. Hall, G. S. Watson, and J. Cabrera. Kernel density estimation with spherical data. *Biometrika*, 74(4):751–762, 1987.

[3] K. V. Mardia and P. E. Jupp. *Directional statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2000.

[4] M. P. Wand and M. C. Jones. *Kernel smoothing*, volume 60 of *Monographs on Statistics and Applied Probability*. Chapman and Hall Ltd., London, 1995.

[5] L. Zhao and C. Wu. Central limit theorem for integrated square error of kernel estimators of spherical density. *Sci. China Ser. A*, 44(4):474–483, 2001.