Available online at www.sciencedirect.com
SCIENCE DIRECT®
Statistics & Probability Letters 73 (2005) 297–304
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Kernel density estimation on Riemannian manifolds

Bruno Pelletier*

Laboratoire de Mathématiques Appliquées, Université du Havre, 25 rue Philippe Lebon, 76063 Le Havre, France

Received 9 March 2004; received in revised form 22 February 2005
Available online 26 April 2005

Abstract

The estimation of the underlying probability density of n i.i.d. random objects on a compact Riemannian manifold without boundary is considered. The proposed methodology adapts the technique of kernel density estimation on Euclidean sample spaces to this nonEuclidean setting. Under sufficient regularity assumptions on the underlying density, L^2 convergence rates are obtained.

Keywords: Nonparametric density estimation; Kernel density estimation; Riemannian manifolds; L^2 convergence

1. Introduction

The situation where the sample space is not Euclidean, but has the structure of a differentiable manifold, may be encountered in numerous fields of science. The case where the sample space is the circle S^1 or the sphere S^2 has been extensively studied, and a great deal of concrete examples is provided by the literature on axial and directional statistics. A survey of statistical methodologies dealing with this kind of data may be found in Jupp and Mardia (1989), Mardia (1972), Watson (1983).

In this paper, we discuss the estimation of a probability density on a Riemannian manifold. The proposed methodology adapts the technique of kernel density estimation on Euclidean sample spaces to this nonEuclidean setting. The manifold is assumed compact without boundary and, to the best of our knowledge, kernel density estimation on this large class of manifolds has not been

*Tel.: +33 2 32 74 43 57; fax: +33 2 32 74 40 24.
E-mail address: Bruno.Pelletier@univ-lehavre.fr.

0167-7152/$ - see front matter
doi:10.1016/j.spl.2005.04.004

studied to date. Density estimation on the circle using trigonometric Fourier series is considered in Devroye and Gyorfi (1985). The generalization of estimation with Fourier series to the case of a compact Riemannian manifold without boundary is developed in Hendriks (1990), where the theory builds upon the eigenfunctions of the Laplace–Beltrami operator on the manifold. Related work on nonparametric deconvolution density estimation on the sphere S^2 may be found in Healy and Kim (1996), Healy et al. (1998), Hendriks (2003). In Hendriks et al. (1993) and Lee and Ruymgaart (1996) the authors consider density and curve estimation on compact smooth submanifolds of a Euclidean space using caps, i.e., intersections of the manifold with closed balls in the ambient Euclidean space. Kernel methods for nonparametric density estimation for axial or directional data are studied in Hall et al. (1987), Fischer et al. (1993), where the kernels proposed by the authors are normalized functions of the scalar product of the evaluation point x and the observation X_i. Classical models for spherical data such as the von Mises distribution on the circle or rotationally symmetric distribution (Watson, 1983) may be expressed as functions of a scalar product $x^t\mu$, for $x, \mu \in S^d$, which is none other than the cosine of the angle between x and μ, showing that they may also be expressed as functions of the geodesic distance on S^d.

The density estimator discussed in this paper is based on kernels that are functions of the Riemannian geodesic distance on the manifold, and its expression is consistent with the expressions of kernel density estimators in the Euclidean case. This estimator has been used recently for image analysis (Lee et al., 2004). It is shown that the appealing idea of centering a small ''mountain'' on the observations, as mentioned in Van der Vaart (1998), is preserved, in the sense that each observation is an intrinsic mean of its associated kernel, provided that the bandwidth be small enough. The estimator and its first properties are formulated in Section 2. Consistency is studied in Section 3. Under sufficient regularity assumptions on the underlying density, L^2 convergence rates are obtained. For materials on differential geometry, we refer to Boothby (1975), Kobayashi and Nomizu (1969), Chavel (1993), Willmore (1993), Hebey (1997).

2. Definition and first properties

Let (M, g) be a compact Riemannian manifold without boundary of dimension d. We shall assume that (M, g) is complete, i.e., (M, d_g) is a complete metric space, where d_g denotes the Riemannian distance.

Let X be a random object on M, i.e., a measurable map on a probability space $(\Omega, \mathscr{A}, P)$ taking values in $(M, \mathscr{B})$, where $\mathscr{B}$ denotes the Borel σ-field of M. We shall assume that the image measure of P by X is absolutely continuous with respect to the Riemannian volume measure, admitting an a.s. continuous density f on M. The Riemannian volume measure will be denoted by v_g.

Let $X_1, \ldots, X_n$ be i.i.d. random objects on M with density f. Let $K : \mathbb{R}_+ \to \mathbb{R}$ be a nonnegative map such that:

(i) $\int_{\mathbb{R}^d} K(\|x\|)\, d\lambda(x) = 1$, (ii) $\int_{\mathbb{R}^d} xK(\|x\|)\, d\lambda(x) = 0$, (iii) $\int_{\mathbb{R}^d} \|x\|^2 K(\|x\|)\, d\lambda(x) < \infty$, (iv) $\operatorname{supp} K = [0; 1]$, (v) $\sup K(x) = K(0)$,

where λ denotes the Lebesgue measure of $\mathbb{R}^d$. Hence the map $\mathbb{R}^d \ni x \to K(\|x\|) \in \mathbb{R}$ is an isotropic kernel on $\mathbb{R}^d$ supported by the closed unit ball.