Kernel density estimation on Riemannian manifolds

Université du Havre
Laboratoire de Mathématiques Appliquées
25 rue Philippe Lebon, 76063 Le Havre, France
Phone : (+33)2 32 74 43 57
Fax (+33)2 32 74 40 24

Bruno.Pelletier@univ-lehavre.fr

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 non-Euclidean setting. Under sufficient regularity assumptions on the underlying density, L^2 convergence rates are obtained.

Index Terms **— 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 non-Euclidean 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 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 $(\Omega, \mathcal{A}, P)$ taking values in $(M, \mathcal{B})$, where $\mathcal{B}$ 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 $X_1, ..., X_n$ be i.i.d. random objects on M with density f. Let $K : \mathbb{R}_+ \to \mathbb{R}$ be a non-negative 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) $suppK = [0; 1]$, 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.

Let p and q be two points of M. Let $\theta_p(q)$ be the volume density function on M, roughly defined by (Besse, 1978, p. 154):

$$\theta_p : q \mapsto \theta_p(q) = \frac{\mu_{exp_p^* g}}{\mu_{g_p}}(exp_p^{-1}(q)),$$

i.e., the quotient of the canonical measure of the Riemannian metric $exp_p^* g$ on $T_p(M)$ (pullback of g by the map exp_p) by the Lebesgue measure of the Euclidean structure g_p on $T_p(M)$. The volume density function is certainly defined for q in a neighborhood of p. In fact, it may defined globally by using Jacobi fields (Willmore, 1993, p. 219). In terms of geodesic normal coordinates at p, $\theta_p(q)$ equals the square-root of the determinant of the metric g expressed in these coordinates at $exp_p^{-1}(q)$. For p and q in a normal neighborhood U of M, we have $\theta_p(q) = \theta_q(p)$ (Willmore, 1993, p. 221).

We define the density estimator of f as the map $f_{n,K} : M \to \mathbb{R}$ which, to each $p \in M$, associates the value $f_{n,K}(p)$ defined by

$$f_{n,K}(p) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r^d} \frac{1}{\theta_{X_i}(p)} K\left(\frac{d_g(p, X_i)}{r}\right). \tag{1}$$

Remark 2.1 Let M be $\mathbb{R}^d$, with its usual Euclidean metric. Then $\theta_p(q) = 1$ for all $p, q \in M$ and $f_{n,K}(p)$ may be written as $f_{n,K}(p) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{r^d} K\left(\frac{\|p-X_i\|}{r}\right)$ So the expression of $f_{n,K}$ is consistent with the expressions of kernel density estimators in the Euclidean case.
We impose the following condition on the bandwidth:

$$r \leq r_0, \tag{2}$$

for some fixed r_0 such that $0 < r_0 < inj_g(M)$, where $inj_g(M)$ denotes the injectivity radius of M (Chavel, 1993, p. 108). Since M is compact, $inj_g(M)$ is strictly positive by the theorem of Whitehead. Hence the set of allowable values for the bandwidth is not the null set. The condition $r < inj_g(M)$

guarantees for each $p \in M$ the existence of a normal coordinate neighborhood at p containing $B_M(p,r)$, the ball in M centered at p and of radius r. Imposing that $r < r_0$ for some strictly positive $r_0 < inj_g(M)$ allows to deal with compact balls in the proof of Lemma 3.3. This is not restrictive since r_0 may be chosen close to $inj_g(M)$, but above all because asymptotics are concerned with sequences r_n decreasing with the number of observations.

Contrary to the Fourier series based density estimator in Hendriks (1990), the estimator defined by Eq. (1) is a probability density on M. More precisely, $f_{n,K}$ is a nonnegative function on M by assumption on K. Furthermore, let $p_1, ..., p_n$ be a realization of $X_1, ..., X_n$. Then

$$\begin{aligned}
\int_M f_{n,K}(p) dv_g(p) &= \int_M \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\theta_{p_i}(p)} K\left(\frac{d_g(p,p_i)}{r}\right) dv_g(p) && (3)\\
&= \int_M \frac{1}{r^d} \frac{1}{\theta_{p_1}(p)} K\left(\frac{d_g(p,p_1)}{r}\right) dv_g(p) && (4)\\
&= \int_{B_M(p_1,r)} \frac{1}{r^d} \frac{1}{\theta_{p_1}(p)} K\left(\frac{d_g(p,p_1)}{r}\right) dv_g(p). && (5)
\end{aligned}$$

Let $(\mathcal{U}_{p_1}, \exp_{p_1}^{-1})$ be the exponential chart at p_1, with normal coordinates $x^1, ..., x^d$. Under the above condition on the bandwidth, $B_M(p_1, r) \subset \mathcal{U}_{p_1}$. Recall that the integral of a continuous function f with compact support included in the domain of a chart (U, φ) takes the following expression in local coordinates $x^1, ..., x^d$:

$$\int_M f(p) dv_g(p) = \int_{\varphi(U)} |g(x)|^{\frac{1}{2}} (f \circ \varphi)(x) dx, \tag{6}$$

where $|g(x)|$ is the determinant of the components of g expressed in the local coordinates $x^1, ..., x^d$. Let $B(r)$ be the ball of radius r in $\mathcal{U}_{p_1}$, i.e., $B(r) = exp_{p_1}^{-1}(B_M(p_1, r))$. We have:

$$\begin{aligned}
\int_M f_{n,K}(p) dv_g(p) &= \int_{B(r)} \frac{1}{r^d} \frac{1}{\theta_{p_1}(exp_{p_1}(x))} K\left(\frac{\|x\|^2}{r}\right) |g(x)|^{\frac{1}{2}} dx && (7)\\
&= \int_{B(r)} \frac{1}{r^d} K\left(\frac{\|x\|^2}{r}\right) dx && (8)\\
&= 1. && (9)
\end{aligned}$$

The kernels involved in kernel density estimators in the Euclidean case are centered on the observations (Akaike, 1954; Parzen, 1962; Rosenblatt, 1956). Provided the bandwidth is small enough, this property is preserved by the estimator defined by Eq. (1) in the following sense. In fact, the concept of mean is to be replaced with the one of intrinsic mean. An intrinsic mean of a probability measure μ on M may be defined as being a minimizer of the energy functional

$$F(p) = \int_M d_g^2(q,p)d\mu(q). \tag{10}$$

This definition is used in Bhattacharya and Patrangenaru (2002, 2003) for instance, while in (Corcuera and Kendall, 1999; Oller and Corcuera, 1995), the authors mainly consider those points $p \in M$ satisfying the criticality condition:

$$\int_{T_p(M)} u\tilde{\mu}(du) = 0, \tag{11}$$

where $\tilde{\mu}$ is a probability measure on $T_p(M)$ which is mapped onto μ by the exponential map at p. Naturally, being a critical point is a weaker concept than being an intrinsic mean. Intrinsic means are also named centers of mass (Kobayashi and Nomizu, 1969; Karcher, 1977; Emery and Mokobodzki, 1991) or Riemannian barycenters (Oller and Corcuera, 1995) as well as Karcher means (Le, 1998) and Frechet expectations (Pennec and Ayache, 1998).

Let G_p^2M be the set of 2-dimensional subspaces of $T_p(M)$. We denote by $\mathcal{K} : \cup_{p\in M} G_p^2M \to \mathbb{R}$ the sectional curvature of M (Kobayashi and Nomizu, 1969, Vol. I p. 202).

Proposition 2.2 *Let q be an arbitrary point of M. Let $\delta = \sup \mathcal{K}$ be the supremum of the sectional curvatures in M. Let μ be a probability measure on M, absolutely continuous w.r.t. the Riemannian volume measure, and with a density $f_q(p)$ defined by*

$$f_q(p) = \frac{1}{r^d}\frac{1}{\theta_q(p)} K\left(\frac{d_g(p,q)}{r}\right),$$

where $r < \min\{\frac{inj_g(M)}{2}, \frac{\pi}{4\sqrt{\delta}}\}$, and where we set $\frac{\pi}{4\sqrt{\delta}} = +\infty$ when $\delta \leq 0$. Then q is an intrinsic mean of μ.

The proof uses the following two propositions, extracted from (Chavel, 1993, p. 337) who follows Karcher (1977) in a study of the center of mass of a probability measure on a Riemannian manifold.

Proposition 2.3 *Let M be a complete Riemannian manifold. Assume $\mathcal{K} < \delta$. Let B be a weakly-convex subset of M with $diam(B) < \pi/2\sqrt{\delta}$. Then the energy functional F has a unique minimum $\overline{p}$ in B.*

Proposition 2.4 *Let M be a complete Riemannian manifold. Assume $\mathcal{K} < \delta$. Let $B = B_M(p_0, r)$ where $r < \min\{\frac{inj_g(M)}{2}, \frac{\pi}{4\sqrt{\delta}}\}$. Then for $p \in B$ we have*

$$|\mathrm{grad} F|(p) \geq d_g(p, \overline{p})\mathbf{C}(r, \delta),$$

where $\mathbf{C}(r, \delta)$ is a strictly positive constant depending on r and on δ.

Proof of Proposition 2.2:
By considering the exponential chart $(\mathcal{U}_q, \exp_q^{-1})$ at q with normal coordinates $x^1, ..., x^d$, it is easily seen that

$$\int_{T_q(M)} x\tilde{\mu}(dx) = 0.$$

Thus q is a critical point of the energy functional F. Since $r < \min\{\frac{inj_g(M)}{2}, \frac{\pi}{4\sqrt{\delta}}\}$, we also have $r < \min\{\frac{inj_g(M)}{2}, \frac{\pi}{2\sqrt{\delta}}\}$, and under this condition, $B_M(q, r)$ is strongly convex. Thus by Proposition 2.3 and Proposition 2.4, q is the unique minimizer of F. $\square$

3 Consistency

We now turn on to the consistency of $f_{n,K}$. Convergence is considered in $L^2(M)$.

Theorem 3.1 *Suppose f is a two-times differentiable probability density on M with bounded second covariant derivative. Let $f_{n,K}$ be the density estimator defined by Eq. (1) with the bandwidth r satisfying the condition of Eq. (2). Then there exists a constant C_f such that*

$$\mathbb{E}_f \|f_{n,K} - f\|^2_{L^2(M)} \leq C_f \left(\frac{1}{nr^d} + r^4 \right).$$

Consequently, for $r \sim n^{-\frac{1}{d+4}}$ we have $\mathbb{E}_f \|f_{n,K} - f\|^2_{L^2(M)} = O(n^{-\frac{4}{d+4}})$.

The proof uses the usual decomposition of $\mathbb{E}_f \|f_{n,K} - f\|^2_{L^2(M)}$ in terms of integrated squared bias and variance (Van der Vaart, 1998), the upper bounds of which are stated in the next two lemmas.

Lemma 3.2 *Let f be a probability density on M and $f_{n,K}$ its estimator, both of which fulfill conditions in Theorem 3.1. Then there exists a constant C_b such that*

$$\int_M (\mathbb{E}_f f_{n,K}(p) - f(p))^2 \, dv_g(p) \leq C_b r^4.$$

Proof:
The pointwise bias $b(p) = \mathbb{E}_f f_{n,K}(p) - f(p)$ may be expressed as

$$\begin{aligned} b(p) &= \int_{B_M(p,r)} \frac{1}{r^d} K\left(\frac{d_g(q,p)}{r}\right)\left(\frac{1}{\theta_q(p)} f(q) - \frac{1}{\theta_p(q)} f(p)\right) dv_q(q) \\ &= \int_{B_M(p,r)} \frac{1}{r^d} \frac{1}{\theta_p(q)} K\left(\frac{d_g(q,p)}{r}\right) (f(q) - f(p))\, dv_g(q). \end{aligned}$$

We shall now take a covariant Taylor expansion of $f(q)$ around p. The Einstein summation convention is used in the following. In geodesic normal coordinates $x^1, ..., x^d$ at p, we have

$$f(x) = f(0) + (\nabla f(x))_i (0) x^i + R_2(p,x),$$

where $R_2(p,x)$ is the remainder. Since f has bounded second order covariant derivative by assumption, the remainder is bounded above by some constant depending on p, which in turn is uniformly bounded in p since M is compact. Consequently, there exists a constant C_R such that for all $p \in M$ and $x \in B(r) \subset T_p(M)$, $|R_2(p,x)| \le C_R \|x\|^2$. Then it follows that

$$b(p) = \int_{B(r)} \frac{1}{r^d} K\left(\frac{\|x\|}{r}\right) R_2(p,x) dx,$$

since $K(\|x\|)$ has a null mean (on $\mathbb{R}^d$) by assumption. Then for all $p \in M$

$$\begin{aligned} |b(p)| &\le \int_{B(r)} \frac{1}{r^d} K\left(\frac{\|x\|}{r}\right) C_R \|x\|^2 dx \\ &= \left(\int_{B(1)} \|y\|^2 K(\|y\|)\, dy\right) C_R r^2, \end{aligned}$$

from which it follows that

$$\int_M b^2(p) dv_g(p) \le C_R^2 \left(\int_{B(1)} \|y\|^2 K(\|y\|)\, dy\right)^2 Vol(M) r^4,$$

where $Vol(M)$ is the volume of M defined by $Vol(M) = \int_M dv_g(p)$. $\square$

Lemma 3.3 *Let f be a probability density on M and $f_{n,K}$ its estimator, both of which fulfill conditions in Theorem 3.1. Then there exists a constant C_v such that*

$$\int_M \mathrm{Var}_f f_{n,K}(p) dv_g(p) \le C_v \frac{1}{nr^d}.$$

Proof:

$$\begin{aligned}\mathrm{Var}_f f_{n,K}(p) &\leq \frac{1}{nr^{2d}}\mathbb{E}_f \frac{1}{\theta^2_{X_1}(p)} K^2\left(\frac{d_g(p,X_1)}{r}\right)\\ &= \frac{1}{nr^{2d}}\int_M \frac{1}{\theta^2_q(p)}K^2\left(\frac{d_g(p,q)}{r}\right) f(q)dv_g(q).\end{aligned}$$

Integrating both sides over M, and denoting $\int_M \mathrm{Var}_f f_{n,K}(p)dv_g(p)$ by I_V yields

$$\begin{aligned} I_V &\leq \int_M \frac{1}{nr^{2d}}\int_M \frac{1}{\theta^2_q(p)}K^2\left(\frac{d_g(p,q)}{r}\right) f(q)dv_g(q)dv_g(p)\\ &= \frac{1}{nr^{2d}}\int_M f(q)\int_M \frac{1}{\theta^2_q(p)}K^2\left(\frac{d_g(p,q)}{r}\right) dv_g(p)dv_g(q).\end{aligned}$$

The integral over p is bounded above by $K^2(0)\int_{B_q(M,r)}\frac{1}{\theta^2_q(p)}dv_g(p)$. Let $C_g(q) = \sup_{B_M(q,r_0)}\theta_q^{-1}(p)$ and let $C_g = \sup_M C_g(q)$. Now this integral is bounded above by $K^2(0)C_g\int_{B_q(M,r)}\frac{1}{\theta_q(p)}dv_g(p) = K^2(0)C_g r^d\omega_d$, where ω_d is the volume of the unit d-dimensional Euclidean sphere. It then follows that

$$\int_M \mathrm{Var}_f f_{n,K}(p)dv_g(p) \leq C_g\omega_d K^2(0)\frac{1}{nr^d},$$

which completes the proof of the lemma.□

Proof of Theorem 3.1:
Write

$$\mathbb{E}_f\|f_{n,K}-f\|^2_{L^2(M)} = \int_M \left(\mathbb{E}_f f_{n,K}(p) - f(p)\right)^2 dv_g(p) + \int_M \mathrm{Var}_f f_{n,K}(p)dv_g(p),$$

and use the above two lemmas for the upper bound. The last assertion is immediate. □

4 Concluding Remarks

An inspection of the proof in Lemma 3.2 shows that the rate obtained in Theorem 3.1 could be improved under additional regularity assumptions on the density f together with nullity conditions on the moments of the kernel up to a given order. The advanced argument is the same as Van der Vaart (1998) for kernel density estimation on the real line. More precisely, suppose f is s times differentiable with bounded s-th covariant derivative. Furthermore, suppose that K is such that $\int_{\mathbb{R}^d} x^I K(\|x\|)dx = 0$, for all multi-indices I of length $|I| < s$ and such that $\int_{\mathbb{R}^d} \|x\|^s K(\|x\|)dx < \infty$, where we have let $x^I = (x^1)^{i_1}...(x^d)^{i_d}$ whenever $I = (i_1, ..., i_d)$. Then the upper bound in Lemma 3.2 becomes at the order of r^{2s}, and the rate in Theorem 3.1 becomes $O(n^{-\frac{2s}{d+2s}})$, for $r \sim n^{-\frac{1}{d+2s}}$, which equals the rate obtained in Hendriks (1990) for density estimation using Fourier series.

Acknowledgements

The author is grateful to the referee for valuable comments. This work was supported by the Department of Mathematics of the University of Le Havre, France, and by the National Science Foundation under Grant No. OCE-0417748.

References

Akaike, H., 1954. An approximation to the density function. Ann. Inst. Statist. Math. 6, 127–132.

Besse, A., 1978. Manifolds all of those geodesics are closed. Ergebnisse der Mathematik und ihrer Grenzgebiete 93. Springer.

Bhattacharya, R., Patrangenaru, V., 2002. Nonparametric estimation of location and dispersion on riemannian manifolds. Journal of Statistical Planning and Inference 108, 23–35.

Bhattacharya, R., Patrangenaru, V., 2003. Large sample theory of intrinsic and extrinsic sample means on manifold (part i). Annals of Statistics 31 (1), 1–29.

Boothby, W., 1975. An Introduction to Differentiable Manifolds and Riemannian Geometry. Academic Press.

Chavel, I., 1993. Riemannian geometry: a modern introduction. Vol. 108 of Cambridge Tracts in Mathematics. Cambridge University Press.

Corcuera, J., Kendall, W., 1999. Riemannian barycentres and geodesic convexity. Mathematical Proceedings of the Cambridge Philosophical Society 127 (2), 253–269.

Devroye, L., Gyorfi, L., 1985. Nonparametric Density Estimation. The L_1-View. Wiley, New York.

Emery, M., Mokobodzki, G., 1991. Sur le barycentre d'une probabilité dans une variété. In: Séminaire de Probabilité, XXV. Vol. 1485 of Lecture Notes in Mathematics. Springer, Berlin, pp. 220–233.

Fischer, N., Lewis, T., Embleton, B., 1993. Statistical Analysis of Spherical Data. Cambridge University Press, Cambridge.

Hall, P., Watson, G., Cabrera, J., 1987. Kernel density estimation with spherical data. Biometrika 74, 751–762.

Healy, D., Hendriks, H., Kim, P., 1998. Spherical deconvolution. J. Multivariate Anal. 67, 1–22.

Healy, D., Kim, P., 1996. An empirical bayes approach to directional data and efficient computation on the sphere. Ann. Statis. 24, 232–254.

Hebey, E., 1997. Introduction à l'analyse non linéaire sur les variétés. Diderot, Paris - New York.

Hendriks, H., 1990. Nonparametric estimation of a probability density on a riemannian manifold using fourier expansions. Ann. Statist. 18, 832–849.

Hendriks, H., 2003. Application of fast spherical fourier transform to density estimation. J. Multivariate Anal. 84, 209–221.

Hendriks, H., Janssen, J., Ruymgaart, F., 1993. Strong uniform convergence of density estimators on compact euclidean manifolds. Statis. Probab. Letters 16, 305–311.

Jupp, P., Mardia, K., 1989. A unified view of the theory of directional statistics 1975-1988. Internat. Statist. Rev. 57, 261–294.

Karcher, H., 1977. Riemannian center of mass and mollifier smoothing. Comm. Pure Appl. Math. 30, 509–554.

Kobayashi, S., Nomizu, K., 1969. Foundations of Differential Geometry. Vol. 1,2. Wiley.

Le, H., 1998. On the consistency of procrustean mean shapes. Adv. in Appl. Probab. 30, 53–63.

Lee, J., Paige, R., Patrangenaru, V., Ruymgaart, F., 2004. Nonparametric density estimation on homogeneous spaces in high level image analysis. In: Aykroyd, R., Barber, S., Mardia, K. (Eds.), Bioinformatics, Images, and Wavelets. pp. 37–40.

Lee, J., Ruymgaart, F., 1996. Nonparametric curve estimation on stiefel manifolds. Nonparam. Statistics 6, 57–68.

Mardia, K., 1972. Statistics of Directional Data. Academic Press.

Oller, J., Corcuera, J., 1995. Intrinsic analysis of statistical estimation. Annals of Statistics 23 (5), 1562–1581.

Parzen, E., 1962. On the estimation of a probability density function and the mode. Ann. Math. Statist. 33, 1065–1076.

Pennec, X., Ayache, N., 1998. Uniform distribution, distance and expectation problems for geometric features processing. J. Math. Imaging Vision 9 (1), 49–67.

Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. Ann. Math. Statist. 27, 832–837.

Van der Vaart, A., 1998. Asymptotic Statistics. Cambridge University Press.

Watson, G., 1983. Statistics on Spheres. Wiley.

Willmore, T., 1993. Riemannian geometry. Oxford University Press.