J Math Imaging Vis (2009) 34: 235–239
DOI 10.1007/s10851-009-0145-2

Kernel Density Estimation on Riemannian Manifolds: Asymptotic Results

Guillermo Henry · Daniela Rodriguez

Published online: 21 February 2009

Abstract The paper concerns the strong uniform consistency and the asymptotic distribution of the kernel density estimator of random objects on a Riemannian manifolds, proposed by Pelletier (Stat. Probab. Lett., 73(3):297–304, 2005). The estimator is illustrated via one example based on a real data.

Keywords Geometry · Nonparametric estimation · Riemannian manifolds · Statistics

1 Introduction

In recent years with the objective to explore the nature of complex nonlinear phenomena, the field of nonparametric inference has increased attention. The idea of nonparametric inference is to leave the data to show the structure lying beyond them, instead of imposing one. Kernel density estimation is a well-known method for estimating the probability density function of a random sample. However, in many applications, the variables X take values on a Riemannian manifold more than on $\mathbb{R}^d$ and this structure of the variables needs to be taken into account when considering neighborhoods around a fixed point X. Some examples could be found in image analysis, astronomy, geology and other fields, they include distributions on spheres, orthogonal groups and Lie groups in general (see [8, 12]). Research on the statistical analysis of such objects was studied by [4, 13] and more recently [16] and [10]. Nonparametric kernel methods for estimating densities of spherical data have been studied by [9] and [1].

Pelletier [14] proposed a family of nonparametric estimators for the density function based on kernel weight when the variables are random object valued in a Riemannian manifolds. More precisely, let (M, g) be a complete Riemannian manifolds and let us consider $X_1, \dots, X_n$ independent and identically distributed random object on M with density function $f(p)$. The Pelletier's idea was to consider an analogue of a kernel on (M, g) by using a positive function of the geodesic distance on M, which is then normalized by the volume density function of (M, g) to take into account for curvature. These estimators are an average of the weight depending on the distance between X_i and p. The Pelletier's estimators is consistent with the kernel density estimators in the Euclidean case. Pelletier [14] studied L^2 convergence rates, under regularity conditions. The object of this note is to complement Pelletier's results with classical properties such as strong uniform consistency and asymptotic distribution.

This paper is organized as follows. Section 2 contains a brief summary of the Pelletier's proposal. Uniform consistency of the estimators is derived in Sect. 3, while in Sect. 4 the asymptotic distribution is obtained under regular assumptions on the bandwidth sequence. Section 5 presents a example using real data. Proofs are given in the Appendix.

This research was partially supported by Grants X-094 from the Universidad de Buenos Aires, PID 5505 from CONICET and PAV 120 and PICT 21407 from ANPCYT, Argentina.

G. Henry · D. Rodriguez (✉)
Departamento de Matemática, FCEyN, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón I, Buenos Aires C1428EHA, Argentina
e-mail: drodrig@dm.uba.ar

G. Henry
e-mail: ghenry@dm.uba.ar

2 Pelletier's Density Estimator

2.1 Preliminaries

Let (M, g) be a d-dimensional oriented Riemannian manifold without boundary. Denote by d_g the distance induced by g and by $\operatorname{inj}_g M = \inf_{p\in M} \sup\{s \in \mathbb{R} > 0 : B_s(p)$ is a normal ball$\}$ the injectivity radius of (M, g).

Throughout this paper, we will assume that (M, g) is complete, i.e. (M, d_g) is a complete metric space, and that $\operatorname{inj}_g M$ is strictly positive. Some examples of Riemannian manifolds with positive injectivity radius are $\mathbb{R}^d$ with g the canonical metric ($\operatorname{inj}_g \mathbb{R}^d = \infty$), and the d-dimensional sphere S^d with the metric induced by the canonical metric of $\mathbb{R}^d$ ($\operatorname{inj}_g S^d = \pi$). It is also well known that compact Riemannian manifolds have positive injectivity radius. Moreover, complete and simply connected Riemannian manifolds with non positive sectional curvature, have also this property. Some standard results on differential geometry can be seen for instance in [2], [3], [5] and [6].

From now on, we will denote by $B_s(p)$ the normal ball in (M, g) centered at p with radius s. Then, $B_s(0_p) = \exp_p^{-1}(B_s(p))$ is an open neighborhood of 0_p in T_pM, the tangent space of M at p, and so it has a natural structure of differential manifold. We are going to consider the Riemannian metrics g' and g'' in $B_s(0_p)$, where $g' = \exp_p^*(g)$ is the pullback of g by the exponential map and g'' is the canonical metric induced by g_p in $B_s(0_p)$. Let $w \in B_s(0_p)$, and $(\bar{U}, \bar{\psi})$ be a chart of $B_s(0_p)$ such that $w \in \bar{U}$. We note by $\{\partial/\partial\bar{\psi}_1|_w, \ldots, \partial/\partial\bar{\psi}_d|_w\}$ the canonical tangent vectors induced by $(\bar{U}, \psi)$. Consider the matricial function with entries (i, j) given by $g'((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))$. The volumes of the parallelepiped spanned by $\{(\partial/\partial\bar{\psi}_1|_w), \ldots, (\partial/\partial\bar{\psi}_d|_w)\}$ with respect to the metrics g' and g'' are given by $|\det g'((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))|^{1/2}$ and $|\det g''((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))|^{1/2}$ respectively. The quotient between these two volumes is independent of the selected chart. So, given $q \in B_s(p)$, if $w = \exp_p^{-1}(q) \in B_s(0_p)$ we can define the volume density function, $\theta_p(q)$, on (M, g) as

$$\theta_p(q) = \frac{|\det g'((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))|^{1/2}}{|\det g''((\partial/\partial\bar{\psi}_i|_w), (\partial/\partial\bar{\psi}_j|_w))|^{1/2}}$$

for any chart $(\bar{U}, \bar{\psi})$ of $B_s(0_p)$ that contains $w = \exp_p^{-1}(q)$. For instance, if we consider the exponential chart (U, ψ) of (M, g) induced by an orthonormal basis $\{v_1, \ldots, v_d\}$ of T_pM and U a normal neighborhood of p then

$$\theta_p(q) = \left|\det g_q\left(\left.\frac{\partial}{\partial\psi_i}\right|_q, \left.\frac{\partial}{\partial\psi_j}\right|_q\right)\right|^{\frac{1}{2}},$$

where $\frac{\partial}{\partial\psi_i}|_q = D_{\alpha_i(0)} \exp_p(\dot{\alpha}_i(0))$ with $\alpha_i(t) = \exp_p^{-1}(q) + t v_i$ for $q \in U$. Note that the volume density function $\theta_p(q)$ is not defined for all p and q in M, but only for those points such that $d_g(p, q) < \operatorname{inj}_g M$. It is worth noticing that, when M is $\mathbb{R}^d$ with the canonical metric, then $\theta_p(q) = 1$ for all $p, q \in \mathbb{R}^d$.

In [11], we calculated the volume density function θ in the case of the two dimensional sphere of radius 1 and in the case of the cylinder. When, we consider the two dimensional sphere of radius R we obtain that the volume density is

$$\theta_p(q) = R\frac{|\operatorname{sen}(d_g(p, q)/R)|}{d_g(p, q)} \quad \text{if } q \neq p, -p \quad \text{and}$$

$$\theta_p(p) = 1.$$

See also, [3] and [16] for a discussion on the volume density function.

2.2 The estimator

Consider a probability distribution with a density f on a d-dimensional Riemannian manifold (M, g). Let $X_1, \ldots, X_n$ be i.i.d. random object takes values on M with density f. In order to estimate f using observations $X_1, \ldots, X_n$ Pelletier [14] proposed the following kernel estimate:

$$f_n(p) = \frac{1}{nh^d}\sum_{i=1}^{n} \frac{1}{\theta_p(X_i)} K\left(\frac{d_g(p, X_i)}{h}\right) \tag{1}$$

where $K : \mathbb{R} \to \mathbb{R}$ is a non-negative function, $\theta_p(q)$ denotes the volume density function on (M, g) and the bandwidth h is a sequence of real positive numbers such that $\lim_{n\to\infty} h = 0$ and $h < \operatorname{inj}_g M$, for all n. This last requirement on the bandwidth guarantees that (1) is defined for all $p \in M$ (see, [15]).

3 Consistency

Let U be an open set of M, we denote by $C^k(U)$ the set of k times continuously differentiable functions from U to $\mathbb{R}$. As in [14], we assume that the image measure of P by X is absolutely continuous with respect to the Riemannian volume measure ν_g, and we denote by f its density on M with respect to ν_g. In this section we will consider the following set of assumptions to derive the strong consistency results of the estimate $f_n(p)$ defined by Pelletier in [14].

H1. Let M_0 be a compact set on M such that:
  (i) f is a bounded function such that $\inf_{p\in M_0} f(p) = A > 0$.
  (ii) $\inf_{p,q\in M_0} \theta_p(q) = B > 0$.

H2. For any open set U_0 of M_0 such that $M_0 \subset U_0$, f is of class C^2 on U_0.

H3. The sequence h is such that $h \to 0$ and $\frac{nh_n^d}{\log n} \to \infty$ as $n \to \infty$.

H4. $K : \mathbb{R} \to \mathbb{R}$ is a bounded nonnegative Lipschitz function of order one, with compact support $[0, 1]$ satisfying: $\int_{\mathbb{R}^d} K(\|\mathbf{u}\|)d\mathbf{u} = 1$, $\int_{\mathbb{R}^d} \mathbf{u}K(\|\mathbf{u}\|)d\mathbf{u} = \mathbf{0}$ and $0 < \int_{\mathbb{R}^d} \|\mathbf{u}\|^2 K(\|\mathbf{u}\|)d\mathbf{u} < \infty$.

Remark 3.1 The fact that $\theta_p(p) = 1$ for all $p \in M$ guarantees that H1 (ii) holds. Assumption H4 is a standard assumption when dealing kernel estimators.

Theorem 3.2 *Assume that* H1 *to* H4 *holds, then we have that*

$$\sup_{p \in M_0} |f_n(p) - f(p)| \xrightarrow{a.s.} 0.$$

4 Asymptotic Normality

Denote by $V \subset M$ an open neighborhood of p. With the objective to derive the asymptotic distribution of the estimator of f, we will consider two assumptions more.

H5. $f(p) > 0$, $f \in C^2(V)$ and the second derivative is bounded.

H6. The sequence h is such that $h \to 0$, $nh^d \to \infty$ as $n \to \infty$ and there exists $0 \le \beta < \infty$ such that $\sqrt{nh^{d+4}} \to \beta$ as $n \to \infty$.

In the following we will denote by $\mathcal{V}_s$ the ball of radius s in $\mathbb{R}^d$ centered at the origin.

Theorem 4.1 *Assume* H4 *to* H6. *Then we have that*

$$\sqrt{nh^d}(f_n(p) - f(p)) \xrightarrow{\mathcal{D}} \mathcal{N}(b(p), V(p))$$

with

$$b(p) = \frac{\beta}{2} \int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_1^2 \, d\mathbf{u} \sum_{i=1}^{d} \frac{\partial f \circ \psi^{-1}}{\partial u_i \partial u_i}\bigg|_{u=0}$$

and

$$V(p) = f(p) \int_{\mathcal{V}_1} K^2(\|\mathbf{u}\|)d\mathbf{u}$$

where $\mathbf{u} = (u_1, \ldots, u_d)$ *and* $(B_h(p), \psi)$ *some exponential chart induced by an orthonormal basis of* T_pM.

Remark 4.2 Note that the Pelletier's estimator converges at the same rate as the Euclidean kernel estimator.

5 Application

We use the data set from [7] to illustrate the use of kernel density estimator on the two dimensional sphere. The data set is the $n = 107$ sites from specimens of Precambrian volcanics whit measurements of magnetic remanence. The data set contains the variables that corresponds to directional component on a longitude scale and on a latitude scale. The original data set can be founded in the library sm of R statistical package.

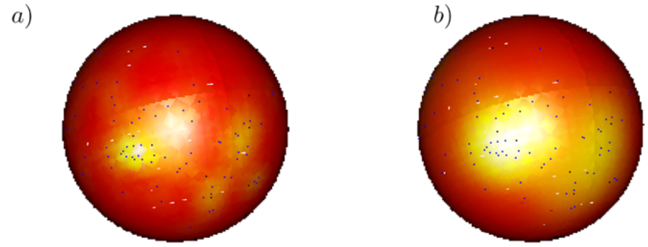

Fig. 1 The nonparametric density estimator using different bandwidth, **a** $h = 1500$ and **b** $h = 3000$

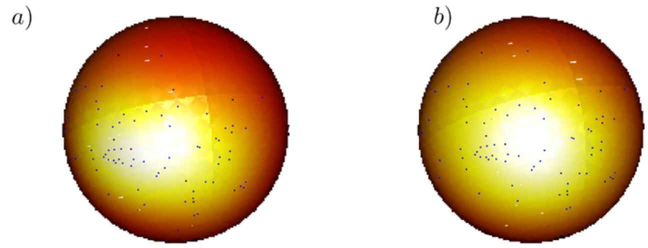

Fig. 2 The nonparametric density estimator using different bandwidth, **a** $h = 5000$ and **b** $h = 7000$

To calculate the estimators the volume density function on the sphere of radius R was taken as in the Sect. 2. We was taken the quadratic kernel $K(t) = (15/16)(1 - t^2)^2 I(|x| < 1)$. The distance d_g between two points p and q in the sphere is

$$d_d(p, q) = R \arccos\left(\frac{\langle p, q \rangle}{\|p\| \|q\|}\right)$$

where $R = 6370$ is the radius of Earth in km and $p = (\cos\alpha \cos\beta, \sin\alpha \cos\beta, \sin\beta)$ with α the longitude and β the latitude.

In order to assess the sensitivity of the results with respect to the bandwidth value, we run the estimation process with different bandwidths. The results are shown in Figs. 1 and 2.

The real data was plotted in blue, in yellow we can be found high levels of concentration of measurements of magnetic remanence while that in red the density estimators is near to 0. As in the Euclidean case large bandwidths produce estimators with small variance but high bias, while small values produce more wiggly estimators. This fact shows the need of the implementation of a method to select the adequate bandwidth for this estimators. However, this require further careful investigation and are beyond the scope of this paper.

Appendix

From now on, we will denote by dv_g the usual volume element induced by g and the orientation of M.

Proof of Theorem 3.2 Using the Theorem 3.2 in [14] and the compactness of M_0 we have that

$$\sup_{p\in M_0} |E(f_n(p)) - f(p)| \le ch^2 \int \|\mathbf{u}\|^2 K(\|\mathbf{u}\|)d\mathbf{u}.$$

Then, in order to obtain strong uniform consistency it suffices to show that $\sup_{p\in M_0} |f_n(p) - E(f_n(p))| \xrightarrow{a.s.} 0$. Let us begin by fixing some notation. Given $p \in M$, denote

$$V_i(p) = \frac{1}{\theta_p(X_i)} K\left(\frac{d_g(p, X_i)}{h}\right) - E\left(\frac{1}{\theta_p(X_i)} K\left(\frac{d_g(p, X_i)}{h}\right)\right),$$

let $S_n(p) = \sum_{i=1}^n V_i$. The fact that $E(V_j) = 0$, the kernel K is bounded and the volume density function satisfies $\theta_p(q) \ge B > 0$ for all $p, q \in M_0$, we have that $|V_i(p)| < A_1$. Then, Bernstein's inequality implies that, for $n > n_0$ and for some positive constants α and $\varepsilon > 0$, we have

$$\sup_{p\in M_0} P\left(\frac{1}{nh^d}|S_n(p)| > \varepsilon\right) \le 2\exp(-nh^d\alpha). \tag{2}$$

On the other hand, since M_0 is a compact set, we can consider a finite collection of balls ($B_i = B_{h^\gamma}(p_i)$) centers at points $p_i \in M_0$ with radius h^γ with $\gamma > 2 + d$, such that $M_0 \subset \bigcup_{i=1}^l B_i$. Then, $l = O(h^{-\gamma})$ and

$$\sup_{p\in M_0} |S_n(p)| \le \max_{1\le j\le l} \sup_{p\in B_j} |S_n(p) - S_n(p_j)| + \max_{1\le j\le l} |S_n(p_j)|. \tag{3}$$

Using that K is a Lipschitz function with Lipschitz constant $\|K\|_L$, straightforward calculation lead to

$$\frac{1}{nh^d}|S_n(p) - S_n(p_j)| < 2B^{-1}\|K\|_L \frac{1}{nh^d} nh^{\gamma-1} = Ch^{\gamma-(d+1)}$$

for all $p \in B_j$, which entails that for n large enough, let us say, for $n > n_1$, we have

$$\max_{1\le j\le l} \sup_{p\in B_j} \frac{1}{nh^d}|S_n(p) - S_n(p_j)| < \varepsilon. \tag{4}$$

Finally, (2), (3) and (4) implies that, for $n > \max\{n_0, n_1\}$

$$P\left(\sup_{p\in M_0} \frac{1}{nh^d}|S_n(p)| > 2\varepsilon\right) \le P\left(\max_{1\le j\le l} |S_n(p_j)| > \varepsilon\right) \le 2ln^{-\alpha\delta_n}$$

with $\delta_n = \frac{nh^d}{\log n}$. By H3 we have that for $n \ge n_2$, $nh^d > 1$, therefore $ln^{-\alpha\delta_n} < Cn^{\gamma/d-\alpha\delta_n}$ for $n \ge n_3$. Since $\delta_n \to \infty$, we have that $\gamma/d - \delta_n\alpha < -2$ for $n \ge n_4$. Hence, for $n \ge \max_{0\le i\le 4} n_i$ and some constant C', we get

$$P\left(\sup_{p\in M_0} \frac{1}{nh^d}|S_n(p)| > 2\varepsilon\right) \le C'n^{-2}$$

which shows that $\sum_{n=1}^\infty P(\sup_{p\in M_0} \frac{1}{nh^d}|S_n(p)| > 2\varepsilon) < \infty$, concluding the proof. □

Proof of Theorem 4.1 Let $S_n(p) = \sum_{i=1}^n V_i(p)$ like in the previous theorem, with $V_i(p) = \frac{1}{\theta_p(X_i)} K(\frac{d_g(p,X_i)}{h}) - E(\frac{1}{\theta_p(X_i)} K(\frac{d_g(p,X_i)}{h}))$. Firstly we note that if we take a Taylor expansion of f around p at order two we get

$$\sqrt{nh^d}(E(f_n(p)) - f(p)) = \frac{\sqrt{nh^d}}{2} \sum_{i,j=1}^d \left[\left.\frac{\partial f\circ\psi^{-1}}{\partial u_i u_j}\right|_{\mathbf{u}=\mathbf{0}} \int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_i u_j d\mathbf{u}\right]h^2 + \sqrt{nh^p}o(h^2)$$

where $\mathbf{u} = (u_1, \dots, u_d)$ and $(B_h(p), \psi)$ is an exponential chart induced by an orthonormal basis of T_pM. Then, the fact that $\int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_i u_j d\mathbf{u} = 0$ and $\int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_i^2 d\mathbf{u} = \int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_j^2 d\mathbf{u}$ if $i \ne j$ implies that

$$\sqrt{nh^d}(E(f_n(p)) - f(p)) \to \frac{\beta}{2} \sum_{i=1}^d \left.\frac{\partial f\circ\psi^{-1}}{\partial u_i \partial u_i}\right|_{\mathbf{u}=\mathbf{0}} \int_{\mathcal{V}_1} K(\|\mathbf{u}\|)u_1^2 d\mathbf{u}.$$

Therefore, it is enough to obtain the asymptotic behavior of $\frac{1}{\sqrt{nh^d}} S_n(p)$. In order to prove this, we will show that $\{V_i(p)\}_i$ satisfies the Lindeberg-Feller Central Limit Theorem.

Pelletier [15] obtained that

$$\frac{1}{nh^d} \operatorname{var}\left(\sum_{i=1}^n V_i(p)\right) = f(p) \int_{\mathcal{V}_1} K^2(\|\mathbf{u}\|)d\mathbf{u} + o(1).$$

Finally, let $\varepsilon > 0$ and we note that

$$h^{-d} E\left(V_1^2(p)\mathbf{1}(V_1(p) > \sqrt{nh^d}\varepsilon)\right) \le h^{-d} E\left(\frac{1}{\theta_p^2(X_1)} K^2\left(\frac{d_g(p, X_1)}{h}\right)\mathbf{1}(V_1(p) > \sqrt{nh^d}\varepsilon)\right)$$

$$+ h^{-d}\left[E\left(\frac{1}{\theta_p(X_1)}K\left(\frac{d_g(p,X_1)}{h}\right)\right)\right]^2$$
$$= A_n + B_n.$$

Since

$$E\left(\frac{1}{\theta_p(X_1)}K\left(\frac{d_g(p,X_1)}{h}\right)\right)$$
$$\leq h^d f(p) + Ch^{2+d}\int_{\mathcal{V}_1}\|\mathbf{u}\|^2 K(\|\mathbf{u}\|)d\mathbf{u}$$

we have that $B_n \to 0$ as $n \to \infty$. Also, exists n_0 such that for all $n \geq n_0$ $|V_1(p)| \leq \frac{1}{\theta_p(X_1)}K(\frac{d_g(p,X_1)}{h}) + \frac{\varepsilon}{2}\sqrt{nh^d}$ then,

$$A_n \leq h^{-d}E\left(\frac{1}{\theta_p^2(X_1)}K^2\left(\frac{d_g(p,X_1)}{h}\right)\mathbf{1}(\theta_p(X_1)^{-1}\right.$$
$$\left.\times |K(d_g(p,X_1)/h)| > \sqrt{nh^d}\varepsilon)\right).$$

Therefore, the fact that $h^{-d}E(\frac{1}{\theta_p^2(q)}K^2(\frac{d_g(p,q)}{h})) = f(p)\int_{\mathcal{V}_1} K^2(\|\mathbf{u}\|)d\mathbf{u} + o(1) < \infty$ implies that $A_n \to 0$ as $n \to \infty$ and $\sum_{i=1}^n E(\frac{1}{nh^d}V_i^2(p)I(X_i)) \to 0$ as $n \to \infty$. Then, we conclude that $\frac{1}{\sqrt{nh^d}}S_n(p) \xrightarrow{\mathcal{D}} N(0, f(p)\int K^2(\|\mathbf{u}\|)d\mathbf{u})$. □

References

1. Bai, Z.D., Rao, C., Zhao, L.: Kernel estimators of density function of directional data. J. Multivar. Anal. **27**, 24–39 (1988)
2. Berger, M., Gauduchon, P., Mazet, E.: Le Spectre d' une variété Riemannienne. Springer, Berlin (1971)
3. Besse, A.: Manifolds all of whose Geodesics are Closed. Springer, Berlin (1978)
4. Bhattacharya, R., Patrangenaru, V.: Nonparametric estimation of location and dispersion on Riemannian manifolds. J. Stat. Plan. Inference **108**, 23–35 (2002)
5. Boothby, W.M.: An Introduction to Differentiable Manifolds and Riemannian Geometry. Academic Press, New York (1975)
6. Do Carmo, M.: Geometria Riemaniana, Proyecto Euclides, 2dn edn. IMPA (1988)
7. Fisher, N.I., Lewis, T., Embleton, B.J.J.: Statistical Analysis of Spherical Data. Cambridge University Press, New York (1987)
8. Goh, A., Vidal, R.: Unsupervised Riemannian clustering of probability density functions. In: Lecture Notes in Artificial Intelligence, vol. 5211 (2008)
9. Hall, P., Watson, G.S., Cabrera, J.: Kernel density estimation with spherical data. Biometrika **74**, 751–762 (1987)
10. Hendriks, H., Landsman, Z.: Asymptotic data analysis on manifolds. Ann. Stat. **35**(1), 109–131 (2007)
11. Henry, G., Rodriguez, D.: Robust nonparametric regression on Riemannian manifolds. J. Nonparametr. Stat. (2009, to appear)
12. Joshi, J., Srivastava, A., Jermyn, I.H.: Riemannian analysis of probability density functions with applications in vision. In: Proc. IEEE Computer Vision and Pattern Recognition (2007)
13. Mardia, K.: Statistics of Directional Data. Academic Press, London (1972)
14. Pelletier, B.: Kernel density estimation on Riemannian manifolds. Stat. Probab. Lett. **73**(3), 297–304 (2005)
15. Pelletier, B.: Nonparametric regression estimation on closed Riemannian manifolds. J. Nonparametr. Stat. **18**, 57–67 (2006)
16. Pennec, X.: Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements. J. Math. Imaging Vis. **25**, 127–154 (2006)