Nonparametric regression for Directional Data

Graciela Boente and Ricardo Fraiman

1. Introduction. Directional data analysis has been extensively developped in the last twenty years, and various new techniques have appeared to meet the needs of scientific workers dealing with data when the observations are directions. This is a frequent situation often met in astronomy, biology, geology, medicine and meteorology as noted in Mardia (1975(a)) where some nice examples are developped such as the analysis of a cancer cell data, or the analysis of long-period comet data. Recently, a unified view of the theory of directional data with an extensive bibliography was done by Jupp and Mardia (1989).

Parametric methods for directional data have grown up from the earlier papers by Fisher (1953), Watson and Williams (1956), Stephens (1969), and a series of papers by Mardia (1967, 1969, 1972(a), 1972(b), 1975(a), 1975(b)). Some more recent results can be found in Jupp and Mardia (1980), Chang (1986, 1989), Prentice (1989), Liu and Singh (1988) and the references mentioned therein.

In recent years a great deal of work has also been done in nonparametric regression methods, and these methods are now more and more used in applications in hidrology, meteorology, medicine, economy, biology and all subjects where modelling is required and there is no a clear parametric model provided. Some examples are given in Watson (1964), Cleveland (1979), Gasser et al.(1984), Yakowitz (1987) among others. These smoothers can be classified into three families: a) convolution type methods (kernel, nearest neighbor and nearest neighbor with kernel methods); b) splines estimates; c) Fourier-type estimates (estimates based on finite dimensional Fourier approximation). However, the first two families are the most extensively studied. A review on asymptotic results for these families can be found in Collomb (1981) and Silverman (1985) respectively.

Some work has also been done in nonparametric methods for directional data as mentioned in Jupp and Mardia (1989). A kernel type smoothing algorithm was given by Watson (1985) although most of the work done in nonparametric spherical regression is based on spline methods (see, for instance, Fisher and Lewis (1985), Watson (1985), Jupp and Kent (1987)).

In this paper we are interested in nonparametric regression methods for directional data. We develop some nonparametric estimation methods for a regression curve when data and the true regression curve itself are at the $p-1$ dimensional sphere, $\mathcal{S}_{p-1}$ based on general weight functions.

Since nonparametric regression functions can be viewed, for random carriers, as conditional expectation, our approach to the problem will be to define a natural notion of conditional expectation on the sphere, and then to use a empirical weighted version to estimate this functional.

The paper is divided as follows. In Section 2 conditional expectation for directional data is defined, and it is shown that the usual parametric models are included as particular cases. In Section 3 normalized weighted averages of the responses variables are considered to estimate the regression function. Consistency, strong convergence rates and asymptotic distribution results are given for kernel, nearest neighbor and nearest neighbor with kernel weights. Confidence regions are also provided under some regularity assumptions. Finally in Section 4, the case of dependent data is discused.

2. Conditional expectation for directional data. We will define a notion of conditional expectation for directional data; in order to provide nonparametric regression models and estimates when data are at the $p-1$ dimensional sphere, $\mathcal{S}_{p-1}$.

Let (X, Y) be a random vector with $X \in \mathbf{R}^d$, $Y \in \mathcal{S}_{p-1}$. Denote by $\| \ \|$ the euclidean norm in $\mathbf{R}^p$, by μ the probability measure associated to the vector X and by $r(x) = E(Y|X = x)$.

It is quite natural to define the conditional expectation on the sphere $\mathcal{S}_{p-1}$ of Y given X as the unique function g which minimizes

$$E\|Y - h(X)\|^2 \qquad \text{for} \quad h \in L^2(\mu) \tag{2.1}$$

subject to the constraint $\|h(X)\| = 1$ a.s. (μ); i.e. if $\mathcal{L} = \{h \in L^2(\mu) : \mu\{\|h(X)\| = 1\} = 1\}$

$$g(X) = \arg\min_{h \in \mathcal{L}} E(\|Y - h(X)\|^2) \tag{2.2}$$

However, if $P(r(X)=0)>0$, the function $g(X)$ will not be uniquely defined. More precisely, denote by $A=\{r(X)=0\}$, by A^c the complement of A and by I_B the indicator function of the set B. Since

$$\begin{aligned} E\|Y-h(X)\|^2 &= E\|Y\|^2 - 2E[Y'h(X)] + 1 = E\|Y\|^2 - 2E[r(X)'h(X)] + 1 \\ &= E\|Y\|^2 - 2E[r(X)'h(X)I_{A^c}] + 1 \end{aligned}$$

for $h \in \mathcal{L}$, the Cauchy Schwartz inequality entails that $h(X) = r(X)\, I_{A^c}/\|r(X)I_{A^c}\| + h_1(X)I_A$, where $h_1(X)$ is any function of $\mathcal{L}$, is a minimizer of (2.1).

Thus, if $r(x_0) \neq 0$ any minimizer will satisfy $h(x_0) = r(x_0)/\|r(x_0)\|$. On the other hand, if $r(x_0)=0$ the mean direction will not be identifiable and there is no a natural way to define it.

REMARK 2.1. Note that if the conditional distribution of $Y|X=x$ is the von Mises–Fischer distribution $F(g(x),\kappa)$, $\kappa>0$, $\|g(x)\|=1$, with density $f(y,g(x),\kappa) = a(\kappa)^{-1}\ \exp(\kappa\, g(x)'y)$ with respect to the uniform distribution on the sphere, the conditional expectation defined above is just the mean direction $g(x)$. Moreover, the same holds for the conditional Fisher–Bingham family (see Mardia (1975(a))) and for the models with symmetry introduced by Saw (1978, 1984) (see also Jupp and Mardia (1989)).

3. Nonparametric regression estimation for data in the sphere.

Let (X_i, Y_i) $1 \le i \le n$ be i.i.d. random vectors $Y_i \in \mathcal{S}_{p-1}$, $X_i \in \mathcal{R}^d$. Denote by $\{W_{ni}(x) \quad 1 \le i \le n\}$ a sequence of probability weight functions, $W_{ni}(x) = W_{ni}(x, X_1, \ldots, X_n)$. Possible examples include the usual weight functions such as kernel weights, nearest neighbor and nearest neighbor with kernel weights.

Let $r_n = \sum_{i=1}^n W_{ni}(x)Y_i$, be the usual nonparametric estimate of the regression function $r(x)$.

Briefly the weight functions mentioned above can be described as:

(a) The kernel-type methods, introduced for regression by Nadaraya (1964) and Watson (1964), correspond to

$$W_{ni}(x) = K((X_i - x)/h_n) \Big/ \sum_{j=1}^n K((X_j - x)/h_n), \tag{3.1}$$

where $h = h_n$ is a sequence of real positive numbers and K is a non-negative real function on R^d with $\int K(u)du < \infty$.

(b) For each $1 \leq j \leq n$ define $I_{nj}(x) = \{i : \|X_i - x\| > \|X_j - x\|\}$. Rank the (X_i, Y_i), $1 \leq i \leq n$, according to increasing values of $\|X_i - x\|$ and obtain a vector of indices $(R_1, \ldots, R_n)$ where X_{R_i} is the ith nearest neighbor of x for all i. Let $k = k_n$ be a sequence of positive integers, $1 \leq k \leq n$. A weight function $\{W_{ni}(x)\}$ is called a k-nearest neighbor weight function if $W_{ni}(x) = 0$ for all $i \in I_{nR_k}(x)$.

(c) The nearest neighbor methods studied, for instance, by Stone (1977) and by Devroye (1981, 1982) correspond to the sequence $\{W_{ni}\}$ satisfying

$$W_{ni}(x) = (v_{n,v_i} + \cdots + v_{n,v_i+\lambda_i-1})/\lambda_i \qquad \text{for} \quad 1 \leq i \leq n\,, \tag{3.2}$$

where $v_i = 1 + \#\{j : 1 \leq j \leq n,\ j \neq i, \text{ and } \|X_j - x\| < \|X_i - x\|\}$, $\lambda_i = 1 + \#\{j : 1 \leq j \leq n,\ j \neq i, \text{ and } \|X_j - x\| = \|X_i - x\|\}$ and $\{v_{ni},\ i \geq 1\}$ is a sequence of real numbers such that $v_{n1} \geq v_{n2} \cdots \geq v_{nn} \geq 0$, $v_{ni} = 0$, for $i > n$ and $v_{n1} + \cdots + v_{nn} = 1$. If $v_{ni} = 0$ for $i > k_n$ we obtain a k-NN weight function.

(d) Denote by $H_n = H_n(x) = \|X_{R_k} - x\|$, where $k = k_n$ is a sequence of positive integers, $1 \leq k \leq n$. Then, the nearest neighbor with kernel weights, which were introduced for regression by Collomb (1980) are defined by

$$W_{ni}(x) = K((X_i - x)/H_n) \Big/ \sum_{j=1}^{n} K((X_j - x)/H_n)\,, \tag{3.3}$$

where $K : \mathbf{R}^d \to R$ is a non-negative function on $\mathbf{R}^d$.

If $K(u) = 0$ for $\|u\| > 1$, we also obtain a k-NN weight function. The choice of $K(u) = I_{\|t\| \leq 1}(u)\lambda(V_1)^{-1}$, where $\lambda(V_1)$ is the Lebesgue measure of the unit ball on $\mathbf{R}^d$ leads to the more usual k-nearest neighbor estimates.

We define, in a natural way, a nonparametric regression estimate in the sphere as $g_n(x) = r_n(x)/\|r_n(x)\|$ if $r_n(x) \neq 0$ and as $g_n(x) = e_1$ if $r_n(x) = 0$ with e_1 the first canonical vector in $\mathbf{R}^p$. In this section, we will obtain consistency results for $g_n(x)$ and its asymptotic distribution. Obviously, $g_n(x)$ is the conditional expectation in the sphere with respect to the conditional empirical distribution.

For the sake of notational simplicity, we will assume that $P(A) = 0$; otherwise, the results will be valid changing the statement "for almost all $x(\mu)$" by "for almost all x in $A^c(\mu)$".

Consistency.

Consider the following assumptions:

A1. There exists a sequence $\{c_n : n \geq 1\}$ of real numbers such that $c_n \geq 0$, $c_n \log n \to 0$, $nc_n \to \infty$ as $n \to \infty$, for which $\max\limits_{1 \leq j \leq n} W_{nj}(x, X_1, \ldots, X_n) \leq c_n$ a.s. for almost all $x(\mu)$.

A2. There exists a random variable K_n and a real number $c > 0$ verifying $\sum_{i \in I_{nR_{K_n}}} W_{ni}(x, X_1, \ldots, X_n) \to 0$ as $n \to \infty$ a.s. for almost all $x(\mu)$, and $\sup\limits_n (c_n K_n) < c$ a.s. for almost all $x(\mu)$.

Conditions under which A1 and A2 are fulfilled are given in Boente and Fraiman (1989(a)).

Let $r(x) = (r_1(x) \ldots r_p(x))'$ $r_n(x) = (r_{1n}(x), \ldots, r_{pn}(x))'$.

THEOREM 3.1. Let $\{W_{ni}(x) \quad 1 \leq i \leq n\}$ be a probability weight functions verifying A1 and A2. Then we have that $g_n(x) \to g(x)$ a.s. for almost all $x(\mu)$.

Proof. Since $\|Y\| = 1$ from Lemma 3.1 of Boente and Fraiman (1989(a)) we obtain $r_{jn}(x) \to r_j(x)$ a.s. for almost all $x(\mu)$. Therefore $r_n(x) \to r(x)$ and $\|r_n(x)\| \to \|r(x)\|$ a.s. for almost all $x(\mu)$; which completes the proof since $P(A) = 0$. ∎

REMARK 3.1. For the weight functions corresponding to kernel methods A1 is not fulfilled, however the conclusion of Theorem 3.1 still hold if K and h_n satisfy the assumptions:

(i) $h_n \to 0 \qquad nh_n^d / \log n \to \infty$ as $n \to \infty$

(ii) There exist positive constants, r, c_1, c_2, c_3, and a bounded Borel function H decreasing on $(0, +\infty)$ such that $c_1 H(\|x\|) \leq K(x) \leq c_2 H(\|x\|)$, $c_3 I_{\|x\| \leq r}(x) \leq K(x)$, and $t^d H(t) \to 0$ as $t \to \infty$,

by applying, for instance, Theorem 2 of Greblicki, Krzyzak and Pawlak (1984) in the proof.

Strong convergence rates.

In Boente and Fraiman (1991 Lemma 2.1) strong convergence rates of the conditional distribution function were obtained. In a similar way, strong convergence rates of the regression function can be derived under the following additional assumptions:

A3. The vector X has a density continuous and positive at x.

A4. For each $1 \leq j \leq p$ $r_j(x)$ is a Lipschitz function, i.e., there exists $\delta > 0$ and $c > 0$ such that

$$\|u - x\| < \delta \Longrightarrow |r_j(x) - r_j(u)\| \leq c\|u - x\|.$$

A5. There exists $c > 0$ such that $P(\theta_n^{-1} \sum_{i \in I_{nR_{K_n}}} W_{ni}(x) \leq c) = 1$, where $\theta_n = (c_n \log n)^{1/2}$.

A6. There exists $a_0 > 0$ such that $a_0 < c_n^{1+2/d} n^{2/d} \log n$ for all n.

Since $P(A) = 0$, from the following equality

$$\theta_n^{-1}\left[\|r_n(x)\|^2 - \|r(x)\|^2\right] = \sum_{j=1}^{p} \left[\theta_n^{-1}(r_{j_n}(x) - r_j(x))\right]\left[r_{j_n}(x) + r_j(x)\right]$$

it is easy to derive the following result

THEOREM 3.2. Under A1 to A6 we have that

$$\theta_n^{-1}(g_n(x) - g(x)) = 0(1) \quad \text{a.s.}$$

REMARK 3.2. As noted in Boente and Fraiman (1991) the conclusion of Theorem 3.2 will also hold for kernel weights under A3 and A4 provided that the sequence $\{h_n : n \geq 1\}$ and the kernel K satisfy the conditions given in Remark 3.1 and the following additional conditions: $h_n \theta_n^{-1} \leq A < \infty$ for all n, where $\theta_n = (\log n / n h_n^d)^{1/2}$ and $t^{d+2} H(t)$ is bounded.

Asymptotic Distribution.

In Theorem 3.3 we derive the asymptotic distribution of $g_n(x)$ by reducing the problem to obtain the asymptotic distribution of the classical nonparametric regression estimate, $r_n(x)$.

THEOREM 3.3. Let x be a point such that $\|r(x)\| > 0$. Assume that there exists a sequence of positive numbers θ_n, converging to 0, such that $\theta_n^{-1/2}(r_n(x) - r(x)) \xrightarrow{w} N(\eta, B)$ with B positive definite, then we have that

$$\theta_n^{-1/2}(g_n(x) - g(x)) \xrightarrow{w} N(\eta_1, B_1)$$

where $\eta_1 = \|r(x)\|^{-1} H\eta$, $B_1 = \|r(x)\|^{-2} HBH'$, $H = I - \delta\delta'$, $\delta = g(x) = r(x)/\|r(x)\|$, I is the identity matrix in $\mathbf{R}^{p \times p}$ and $\xrightarrow{w}$ stands for weak convergence.

REMARK 3.3. Clearly, the asymptotic covariance matrix B_1 has rank $p-1$ and corresponds to the covariance matrix of a gaussian vector Z which can be obtained from a gaussian vector U, with mean η and covariance matrix B, through a projection, rescaled by $\|r(x)\|^{-1}$, in the orthogonal complement of $r(x)$.

In order to prove Theorem 3.3 we will need the following Lemma due to Rubin and proved by Anderson (1963).

LEMMA 3.1. Let $F_n(u)$ be the cumulative distribution function of a random matrix U_n. Let V_n be a (matrix-valued) function of U_n, $V_n = f_n(U_n)$ and let $G_n(v)$ be the (induced) distribution of V_n. Suppose $\lim_{n\to\infty} F_n(u) = F(u)$ [in every continuity point of $F(u)$] and suppose for every continuity point u of $f(u)$, $\lim_{n\to\infty} f_n(u_n) = f(u)$ when $\lim_{n\to\infty} u_n = u$. Let $G(v)$ be the distribution of the random matrix $V = f(U)$, where U has the distribution $F(u)$. If the probability of the set of discontinuities of $f(u)$ according to $F(u)$ is 0, then $\lim_{n\to\infty} G_n(v) = G(v)$ [in every continuity point of $G(v)$].

PROOF OF THEOREM 3.3. Denote by $U_n = \theta_n^{-1/2}[r_n(x) - r(x)]$ and by $V_n = \theta_n^{-1/2}[g_n(x) - g(x)]$. We have that $U_n \xrightarrow{w} U$ where U is a random vector with distribution $N(\eta, B)$. Let

$$\begin{aligned} f_n(u) = {} & [u/\|\theta_n^{1/2}u + r(x)\|] - \\ & - \delta[\theta_n^{1/2}\|u\|^2 + 2r(x)'u]/\{\|\theta_n^{1/2}u + r(x)\|[\|r(x)\| + \|\theta_n^{1/2}u + r(x)\|]\} \\ f(u) = {} & [u/\|r(x)\|] - \delta\|r(x)\|^{-2}r(x)'u = \|r(x)\|^{-1}Hu \end{aligned}$$

Then it is easy to see that $V_n = f_n(U_n)$ and that $f_n(u_n) \to f(u)$ as $n \to \infty$ when $u_n \to u$ as $n \to \infty$. Thus, from Lemma 3.1 we have that $V_n \xrightarrow{w} f(U)$ which concludes the proof. ∎

We will now derive an explicit form for the asymptotic bias and covariance matrix of the regression estimates related to kernel, nearest neighbor and nearest neighbor with kernel weights under the following assumptions:

N1. The kernel $K : \mathbf{R}^d \to R$ is bounded, nonnegative, $0 < \int K^2(u)du < \infty$, and $\|u\|^d K(u) \to 0$ as $\|u\| \to \infty$.

N2. There exists $0 \le \beta < \infty$ such that $h_n n^{1/(d+2)} \to \beta$ as $n \to \infty$.

N3. The vector X has a density $f(\cdot)$ which is positive at x.

Denote by $\sigma_j^2(u) = E[(Y_{j1} - r_j(x))^2 | X_1 = u]$, by $\ell_{tj}(u) = E(Y_{j1}Y_{t1} | X_1 = u)$ and by $\Sigma(x)$ the covariance matrix of $Y_1 | X_1 = x$.

N4. (a) r_j verifies a Lipschitz condition of orden one at x and there exists $\lim_{\varepsilon\to 0}[r_j(x+\varepsilon u)-r_j(x)]/\varepsilon = r_j'(x,u)$.

(b) σ_j^2 and ℓ_{tj} are continuous in a neighborhood of x.

N5. The kernel K is twice continuously differentiable and verifies:

(a) $0 < \int |K_1(u)|du < \infty$; $\int K_1^2(u)du < \infty$ and $\|u\|^d K_1(u) \to 0$ as $\|u\| \to \infty$ where $K_1(u) = \sum_{j=1}^d \dfrac{\partial K}{\partial u_j}(u)u_j$.

(b) $\|u\|^{d+1}K_2(u) \to 0$ as $\|u\| \to \infty$ where $K_2(u) = \sum_{i,j} \dfrac{\partial^2 K}{\partial u_i \partial u_j}(u)u_i u_j$ and $u = (u_1, \ldots, u_d)$.

N6. There exists $0 \le \beta < \infty$ such that $k_n^{1/d} n^{1/(d+2)-(1/d)} \to \beta$.

Let $v_{n1} \ge \cdots \ge v_{nn} \ge 0$, $\sum_{i=1}^n v_{ni} = 1$ denote by $\tau_n = \sum_{i=1}^n v_{ni}^2$.

N7. $\lim_{n\to\infty} v_{n1} = 0$ and there exists a sequence of positive integers k_n such that $k_n \to \infty$, $k_n/n \to 0$ as $n \to \infty$ and $k_n v_{n1}$ is bounded and $\sum_{j>k_n} v_{nj} \to 0$ as $n \to \infty$.

N8. $v_{n1}/\tau_n^{1/2} \to 0$ as $n \to \infty$.

N9. $\lim_{n\to\infty} \tau_n^{-1} \sum_{j>k_n} v_{nj} = 0$ and $\lim_{n\to\infty} \tau_n^{-1/2}(k_n/n)^{1/d} = 0$.

From Lemmas 3.2, 3.3 and 3.4 of Boente and Fraiman (1991) we obtain

Proposition 3.1. Let (X_i, Y_i) $1 \le i \le n$ be i.i.d. random vectors, $X_i \in \mathbf{R}^d$, $Y_i \in \mathcal{S}_{p-1}$. Assume that N4 holds.
(a) Under A3, N1, N2 and N3 we have that

$$(nh_n^d)^{1/2}(r_n(x) - r(x)) \xrightarrow{w} N(\eta, B)$$

where

$$\eta_i = \beta^{(d/2)+1} \int r_i'(x,u)K(u)du \tag{3.4}$$

$$B_{ij} = \Sigma_{ij}(x) f(x)^{-1} \int K^2(u)du \tag{3.5}$$

when $W_{ni}(x)$ are the kernel weights given by (3.1).
(b) Under A3, N1, N3, N5 and N6 we have that

$$k_n^{1/2}(r_n(x) - r(x)) \xrightarrow{w} N(\eta, B)$$

where

$$\eta_i = \beta^{(d/2)+1} \int r_i'(x,u)K(u)du[f(x)\lambda(V_1)]^{1/2} \tag{3.6}$$

$$B_{ij} = \Sigma_{i,j}(x)\lambda(V_1) \int K^2(u)du \tag{3.7}$$

when $W_{ni}(x)$, given in (3.3), are the nearest neighbor with kernel weights.
(c) Under N7, N8 and N9 we have that

$$\tau_n^{-1/2}(r_n(x) - r(x)) \xrightarrow{w} N(0, B)$$

with $B = \Sigma(x)$ when $W_{ni}(x)$, given in 3.2, are the nearest neighbor weights.

Confidence regions.

Throughout this section we assume $\eta = 0$.

In order to obtain confidence regions we can proceed in two different ways:
(a) to look for circular regions on the sphere, centered at $g_n(x)$, $\mathcal{R}_n = \{u \in \mathcal{S}_{p-1} : \|u - g_n(x)\| \le C_n\}$ which will bring over some problems of numerical integration to obtain the corresponding percentiles, since the covariance matrix is not necessarily idempotent.
(b) to look for confidence regions which adequately transformed will lead to circular regions for which percentiles can be easily derived.

We will describe briefly the second option.

Let $B_1 = \Gamma\Lambda\Gamma$ where $\Gamma\Gamma' = \Gamma'\Gamma = I$, $\Gamma = (\gamma_1, \ldots, \gamma_p)$ and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{p-1}, 0)$ with $\lambda_1 \ge \cdots \ge \lambda_{p-1}$ and B_1 is given in Theorem 3.3. Denote by $\widetilde{\Gamma} = (\gamma_1, \ldots, \gamma_{p-1})$ $\widetilde{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_{p-1})$ then $\widetilde{\Lambda}^{-1/2}\widetilde{\Gamma}'Z \sim N(0, I_{p-1})$, $P(\gamma_p'Z = 0) = 1$ with $Z \sim N(0, B_1)$.

Thus if $\chi^2_{p-1,\alpha}$ denotes the $(1-\alpha)$-percentile of a chi-square distribution with $(p-1)$ degrees of freedom, the region

$$\mathcal{R} = \{u \in \mathcal{S}_{p-1} \; \|\widetilde{\Lambda}^{-1/2}\widetilde{\Gamma}'(u - g_n(x))\|^2 \le \theta_n \chi^2_{p-1,\alpha}\}$$

is a confidence region for $g(x)$ with asymptotic level $1 - \alpha$.

In practice the matrix B_1 which envolves the conditional covariance matrix of $Y_1|X_1 = x$ and the marginal density $f(x)$ is unknown. However, it can be estimated from the data using the empirical distribution function and estimates of the density function.

4. Some extensions

The dependent case. The way in which these results have been obtained allows us to extend them to the dependent case. More precisely, let $\{(X_t, Y_t)\ t \geq 1\}$ be a stationary α-mixing sequence, $X_t \in \mathbf{R}^d$, $Y_t \in \mathcal{S}_{p-1}$. We will only consider kernel and k-nearest neighbor with kernel weights. From Theorems 4.1 and 4.2 of Boente and Fraiman (1989(b)) and Lemma 2 and an argument similar to that used in Theorem 2 of Boente and Fraiman (1990) we obtain the following result:

THEOREM 4.1.
(i) Under H1 to H4 of Boente and Fraiman (1989(b)) we have that $g_T(x) \rightarrow g(x)$ a.s. for almost all $x(\mu)$ where $r_T(x) = \sum_{t=1}^T W_{Tt}(x) Y_t$, $W_{Tt}(x)$ is given by (3.1) and $g_T(x) = r_T(x)/\|r_T(x)\|$.
(ii) Under H1, H2' to H5' of Boente and Fraiman (1989(b)) we have that $\widehat{g}_T(x) \rightarrow g(x)$ a.s. for almost all $x(\mu)$ where $\widehat{r}_T(x) = \sum_{t=1}^T \widehat{W}_{Tt}(x) Y_t$, $\widehat{W}_{Tt}(x)$ are given by (3.3) and $\widehat{g}_T(x) = \widehat{r}_T(x)/\|\widehat{r}_T(x)\|$.
(iii) Let x be a point such that $\|r(x)\| > 0$. Assume that the mixing coefficients verify $N \sum_{j=N+1}^{\infty} \alpha(j) \rightarrow 0$ as $N \rightarrow \infty$ and that for all $s \geq 1$ the density $f_s(u, v)$ of (X_t, X_{t+s}) is bounded uniformly in s. Moreover, assume that N1 to N4 hold. Then, we have that

$$(T h_T^d)^{1/2} (g_T(x) - g(x)) \xrightarrow{w} N(\eta_1, B_1)$$

with $\eta_1 = \|r(x)\|^{-1} H \eta$, $B_1 = \|r(x)\|^{-2} H B H'$ where $H = I - g(x) g(x)'$, η and B are given by (3.4) and (3.5).
(iv) Let x be a point such that $\|r(x)\| > 0$. Assume that the assumptions of (iii) and N5, N6 hold, then we have that $k_T^{1/2} (\widehat{g}_T(x) - g(x)) \xrightarrow{w} N(\eta_1, B_1)$ where $\eta_1 = \|r(x)\|^{-1} H \eta$, $B_1 = \|r(x)\|^{-2} H B H'$ and η and B are now given by (3.6) and (3.7).

Strong convergence rates can also be obtained.

References

ANDERSON, T.W. (1963). *Asymptotic theory for principal components.* Ann. Math. Statist. **34,** 122-148.

BOENTE, G. and FRAIMAN, R. (1989(a)) *Robust nonparametric regression estimates.* J. Multivariate Anal. **29,** 180-198.

BOENTE, G. and FRAIMAN, R. (1989(b)) *Robust nonparametric regression estimation for dependent observations.* Ann. Statist. **17,** 1242-1256.

BOENTE, G. and FRAIMAN, R. (1990) *Asymptotic distribution of robust estimates for nonparametric models from mixing processes.* Ann. Statist. **18,** 891-906.

BOENTE, G. and FRAIMAN, R. (1991) *Strong uniform convergence rates for some robust equivariant nonparametric regression estimates for mixing processes.* International Statistical Review**59,** ??????.

CHANG, T. (1986). *Spherical regression.* Ann. Statist. **14,** 907-924.

CHANG, T. (1989). *Spherical regression with errors in variables.* Ann. Statist. **17,** 293-306.

CLEVELAND, W. (1979). *Robust locally weighted regression and smoothing scatterplots.* J. Amer. Statist. Assoc. **74,** 829-836.

COLLOMB, G. (1980). *Estimation de la regression par la methode des k points les plus proches avec noyau: Quelques propriétés de convergence ponctuelle.* Statistique nonparamétrique asymptotique. Lecture Notes in Math. **821,** 159-175.

COLLOMB, G. (1981). *Estimation nonparametrique de la regression: Revue bibliographique.* Internat. Statist. Rev. **49,** 75-93.

DEVROYE, L. (1981). *On the almost everywhere convergence of nonparametric regression function estimates.* Ann. Statist. **9,** 1310-1319.

DEVROYE, L. (1982). *Necessary and sufficient conditions for the pointwise convergence of nearest neighbor regression function estimates.* Z. Wahrsch. verw. Gebiete **66,** 441-460.

FISHER, R.A. (1953). *Dispersion on a sphere.* Proc. Roy. Soc. Lond. A, **217,** 295-305.

FISHER, N.E. and LEWIS, T. (1985). *A note on spherical splines.* J. R. Statist. Soc. B **47,** 482-488.

GASSER, TH., MÜLLER, H., KÖHLER, W., MOLINARI, L. and PRADER, A. (1984). *Nonparametric regression analysis of growth curves.* Ann. Statist. **12,** 210-229.

GEORGIEV, A. (1988). *Consistent nonparametric multiple regression: the fixed design case.* J. Multivariate Anal. **25,** 100-110.

GREBLICKI, W., KRZYZÁK, A. and PAWLAK, M. (1984). *Distribution-free pointwise consistency of the kernel regression estimate.* Ann. Statist. **12,** 1570-1575.

JUPP, P.E. and MARDIA, K.V. (1980). *A general correlation coefficient for directional data and related regression problems.* Biometrika **63,** 163-173.

JUPP, P.E. and MARDIA, K.V. (1989). *A unified view of the theory of directional statistics, 1975-1988.* Int. Statist. Rev. **57,** 261-294.

JUPP, P.E. and KENT, J.T. (1987). *Fitting smooth paths to spherical data.* Appl. Statist. **36,** 34-46.

LIU, R.Y. and SINGH, K. (1988). *On the ordering of directional data: Concepts of data depth on circles and spheres.* Unpublished manuscript.

MARDIA, K.V. (1967). *A non-parametric test for the bivariate two sample location problem.* J. R. Statist. Soc. Serie B **29,** 320-342.

MARDIA, K.V. (1969). *On Wheeler and Watson's test on a circle.* Sankhya, Serie A, **31,** 177-190.

MARDIA, K.V. (1972(a)). *Statistics of Directional Data*. London: Academic Press.

MARDIA, K.V. (1972(b)). *A multi-sample uniform score test on a circle and its parametric competitor.* J. R. Statist. Soc. Serie B **34,** 102-113.

MARDIA, K.V. (1975(a)). *Statistics of directional data.* J. R. Statist. Assoc. **37,** 349-393.

MARDIA, K.V. (1975(b)). *Characterizations of directional distributions.* In *Statistical Distributions for Scientific Work* (G.P. Patil, S. Kotz and J.K. Ord, eds.) Vol 1. Boston: D. Reidel.

NADARAYA, E.A. (1964). *On estimating regression.* Theory Probab. Appl. **9,** 141-142.

PRENTICE, M.J. (1989). *Spherical regression on matched pairs of orientation statistics.* J. R. Statist. Soc. Serie B, **51,** 241-248.

RUBIN, H. *The topology of probability measures on a topological space.* Unpublished manuscript.

SAW, J.G. (1978). *A family of distributions on the m-sphere and some hypothesis tests.* Biometrika **65**, 69-73.

SAW, J.G. (1984). *Ultraspherical polynomials and statistics on the m-sphere.* J. Multivariate Anal. **28**, 91-97.

SILVERMAN, B.W. (1985). *Some aspects of the spline smoothing approach to non-parametric regression curve fitting.* J. R. Statist. Soc. Serie B, **47,** 1-52.

STEPHENS, M.A. (1969). *Tests for the von Mises distribution.* Biometrika **56,** 149-160.

STONE, C. (1977). *Consistent nonparametric regression (with discussion).* Ann. Statist. **5,** 595-645.

WATSON, G.S. (1964). *Smooth regression analysis.* Sankhya Ser. A **26,** 359-372.

WATSON, G.S. and WILLIAMS, E.J. (1956). *On the construction of significance tests on the circle and the sphere.* Biometrika **43,** 344-352.

WATSON, G.S. (1985).

YAKOWITZ, S. (1987). *Nearest-neighbor methods for time series analysis.* J. Time Ser. Anal. **8,** 235-247.

Graciela Boente
Departamento de Matemática
Fac. de Ciencias Exactas y Nat.
Ciudad Universitaria, Pabellón N° 1
Buenos Aires 1428, Argentina.

Ricardo Fraiman
Centro de Matemática
Universidad de la República
Eduardo Acevedo 1139
Montevideo, Uruguay.