Analele ştiinţifice ale Universităţii „Al. I. Cuza" Iaşi
Tomul XXIV, f. 2, s. I a, 1978

CONVERGENCE RATES OF $k(n)$-NEAREST NEIGHBOR DENSITY ESTIMATES*)

BY

LEWIS PAKULA

If $X_1, \ldots, X_n$ is an i.i.d. sequence of random variables in $\mathbb{R}^d$ with continuous density $f > 0$, Loftsgaarden and Quesenberry [8] proposed estimating $f(x)$ by

$$\hat{f}_n(x) = k(n)/n \operatorname{Vol}(S_{k(n)}) \tag{1}$$

where $k(n)$ is a sequence of integers satisfying $k(n) \to \infty$, $k(n)/n \to 0$, and $S_{k(n)}$ is the smallest sphere with center at x containing the $k(n)$ values of X_i nearest to x. They showed that $\hat{f}_n(x)$ is a consistent estimate in probability and under mild conditions on f and $k(n)$, Wagner and Devroye [4] showed that in fact $|\hat{f}_n(x) - f(x)| \to 0$ uniformly on $\mathbb{R}^d$ with probability 1. On the basis of some non-rigorous calculations Fukunaga and Hostetler [2] claim that generally $E[\hat{f}_n(x) - f(x)]^2 = O(n^{-4/(d+4)})$ when $k(n) \sim Cn^{4/(d+4)}$ and that this is the optimal choice of $k(n)$ having the form Cn^{α}. In this note we consider instead the error criterion, $D_n = E[\hat{f}_n(x)^{-1} - f(x)^{-1}]^2$ and prove that for $k(n) = n^{2/3}$, $D_n = O(n^{-2/3})$ under some mild explicit conditions on f, independently of the dimension d. We give an example to show that this cannot be improved in general.

Let $\|\cdot\|$ denote the Euclidean norm in $\mathbb{R}^d$ and c the volume of the unit ball. Let $\xi_{k(n)}$ be the $k(n)$-th order statistic of $c\|x - X_1\|^d, \ldots, c\|x - X_n\|^d$, so that $\hat{f}_n(x) = k(n)/n\xi_{k(n)}$. For convenience take $x = 0$. If X has probability distribution P with density f,

$$f(0) = \lim_{r \to 0} P(S(0;r))/\operatorname{Vol}(S(0;r)) = \lim_{r \to 0} (cr^d)^{-1} \int_{S(0;r)} f(x)\,dx =$$

$$= \lim_{t \to 0} t^{-1} \int_{S(0;(t/c)^{1/d})} f(x)\,dx,$$

* This work was partially supported by a 1977 University of Rhode Island Summer Faculty Fellowship.

where $S(\mathbf{0}; r) = \{ y : \| y \| < r \}$. Thus $f(\mathbf{0}) = G'(0)$ where $G(t)$ is the distribution function of $c \| X \|^d$ and $G'(0)$ is the derivative from the right. G has a continuous density $g > 0$ (right continuous at ó) and $f(\mathbf{0}) = g(0)$. We make the assumption

(A) $E \| X \|^{2d} < \infty$ and g has a continuous, bounded derivative in some open interval $(0, \varepsilon)$.

For $0 < \delta < 1$ let $\theta_\delta = H(\delta)$ where H is the inverse of G, and let $\delta_n = k(n)/n$. We have

$$(2) \qquad D_n \leqslant [[E(\xi_{k(n)} \delta_n^{-1} - \theta_{\delta_n} \delta_n^{-1})^2]^{1/2} + | \theta_{\delta_n} \delta_n^{-1} - g(0)^{-1} |]^2.$$

First, the mean value theorem applied to H shows that for δ sufficiently small, $\theta_\delta \leqslant 2\delta g(0)^{-1}$. Since $\delta = \int_0^{\theta_\delta} g = g(0)\,\theta_\delta + \int_0^{\theta_\delta}\int_0^{x} g'(t)\,dt\,dx$ we have

$$(3) \qquad | (\theta_\delta/\delta) - g(0)^{-1} | \leqslant M \theta_\delta^2 / 2g(0)\,\delta \leqslant M\delta / g(0)^2$$

where $M = \sup | g'(t) |$ in a suitable interval.

If $U_1, \ldots, U_n$ are the order statistics of n independent r.v.'s uniformly distributed on $[0, 1]$ then $H(U_{k(n)})$ has the distribution of $\xi_{k(n)}$ and $\theta_{\delta_n} - \xi_{k(n)}$ is distributed as $H(\delta_n) - H(U_{k(n)})$. Now

$$(4) \qquad E[H(\delta_n) - H(U_{k(n)})]^2 = E[(H(\delta_n) - H(U_{k(n)}))^2,\ U_{k(n)} < 1/2] + E[(H(\delta_n) - H(U_{k(n)}))^2,\ U_{k(n)} \geqslant 1/2].$$

For $0 < s, t < 1/2$ we have $|H(t) - H(s)| \leqslant M_1 | t - s |$ where $M_1 = \sup \{ | g(t) |^{-1} : 0 \leqslant t \leqslant H(1/2) \}$. Using the mean and variance of the (Beta) distribution of $U_{k(n)}$ we see that the first term on the right in (4) is smaller than

$$(5) \qquad M_1^2 [\operatorname{var} U_{k(n)} + (EU_{k(n)} - \delta_n)^2] \leqslant M_1^2 [(n - k)\,kn^{-3} + k^{-2} n^{-4}].$$

The second term is smaller that $\int_{1/2}^{1} H(t)^2 \beta(t)\,dt$ for $\delta_n < 1/2$ where $\beta(t)$, the density of $U_{k(n)}$ is decreasing for large n and $t > 1/2$. Since H is increasing this integral is majorized by $\int_{1/2}^{1} H(t)^2\,dt \,.\, P(U_{k(n)} > 1/2)$.

Let $M_2 = \int_{1/2}^{1} H(t)^2\,dt$, then since $E \| X \|^{2d} < \infty$,

$M_2 \leqslant \int_0^1 H(t)^2\,dt = \int_0^\infty x^2 dG < \infty$. Thus for δ_n sufficiently small we get

$P(U_{k(n)} < 1/2) \leqslant 5 \operatorname{Var} U_{k(n)}$ by Chebyshev's inequality. Combining this with (2), (3), (4), and (5) we find that for n large and hence δ_n small, D_n is smaller than the square of

$$(6) \qquad (n/k)\,[(M_1^2 + 5M_2)\,(k)\,(n - k)/n^3 + M_1^2 k^2/n^4]^{1/2} + Mg(0)^{-2}\,(k/n) \leqslant c_1 (k^{-1} - n^{-1})^{1/2} + c_2 (k/n) \leqslant c_1 k^{-1/2} + c_2 k/n$$

after consolidating some constants. Putting $k \sim n^{2/3}$ we have

Theorem. *Under assumption* (A), $D_n = O(n^{-2/3})$ *if* $k \sim n^{2/3}$.

The estimates above are admittedly crude but we have the following

Example. Suppose $G(x) = 1 - e^{-x}$, $x > 0$ so $g(x) = e^{-x}$, $x > 0$. Then

$$D_n = E[(n\xi_k/k) - 1]^2 = \operatorname{Var}(n\xi_k/k) + [E(n\xi_k/k) - 1]^2 =$$

$$= (n^2/k^2)\,(\sum_{n-k+1}^{n} i^{-2}) + [(n/k)\,(\sum_{n-k+1}^{n} i^{-1}) - 1]^2$$

(See D a v i d [1], p. 39).

An integral approximation to the sum in the first term and some manipulation shows that the first term is $k^{-1} + o(k^{-1})$ recalling that $k/n \to 0$ Similarly $[(n/k)(\Sigma i^{-1}) - 1] = -\delta_n^{-1} \log(1 - \delta_n) + o(k^{-1})$ which by Taylor's formula becomes $\delta_n/2 + o(\delta_n) + o(k^{-1})$. Squaring, we see that the second term is $\delta_n^2/4 + o(\delta_n^2) + o(k^{-1})$ so $D_n = 1/k \pm k^2/4n^2 + o(k^2/n^2) + o(1/k)$. Thus for k of the form n^α, $0 < \alpha < 1$, $D_n \geqslant C_n^{-\beta}$ for some $c > 0$, for $\beta = \min(\alpha, 2 - 2\alpha)$ and $\beta < 2/3$ for $\alpha \neq 2/3$.

We now show that $E\hat{f}_n(0)^2 \to f(0)^2$. Let $H(t)$ be as above. For given ε, $0 < \varepsilon < a$, there is a $\delta > 0$ such that $(a - \varepsilon)\,t \leqslant H(t) \leqslant (a + \varepsilon)\,t$ for $0 \leqslant t < \delta$ where $a = f(0)^{-1}$. We denote by β the density of U_k as before. Then

$$E\hat{f}_n^2 = (k/n)^2 \int_0^1 H(t)^{-2} \beta(t)\,dt \leqslant (k/n)^2 \int_0^1 (a - \varepsilon)^{-2} t^{-2} \beta(t)\,dt + (k/n)^2 \int_\delta^1 H(t)^{-2} \beta(t)\,dt.$$

A simple manipulation of beta integrals a shows that the first term approaches $(a - \varepsilon)^2$ as $n \to \infty$ and $k/n \to 0$, while an application of Chebyshev's inequality and the fact that $1/H(x)$ is bounded for $t > \delta$ shows that the second term approaches 0. Similarly

$$E\hat{f}_n^2 \geqslant (k/n)^2 \int_0^1 (a + \varepsilon)^{-2} t^{-2} \beta(t)\,dt + \int_\delta^1 [H(t)^{-2} - (a + \varepsilon)^{-2} t^{-2}]\,\beta(t)\,dt$$

with the first term approaching $(a + \varepsilon)^{-2}$ and the second, 0. The conclusion follows. Notice that no moment assumptions are necessary here.

Now by the Cauchy-Schwartz inequality we have

$$E\,|\hat{f}_n(x) - f(x)| \leqslant \{ E\,[\hat{f}_n(x)^{-1} - f(x)^{-1}]^2\ E\,[\hat{f}_n(x) f(x)]^2 \}^{1/2}$$

If assumption (A) holds, then, combining this with (6) allows us to estimate the speed of convergence of $E\,|\hat{f}_n(x) - f(x)|$ to 0 for a given choice of $k(n)$. For example with $k(n) \sim n^{2/3}$ we get $E\,|\hat{f}_n(x) - f(x)| = O(n^{-1/3})$. For large d this estimate gives a faster convergence than what is obtained by applying the Cauchy-Schwartz inequality to the estimate of $E[\hat{f}_n(x) - f(x)]^2$ stated in [2] with the optimal choice of $k(n)$ stated there.

REFERENCES

1. David H. A. — *Order Statistics*, New Yok, Wiley, 1970.
2. Fukunaga K. and Hostetler L. D. — *Optimization of k-nearest neighbor density estimates*, IEEE Transactions on Information Theory *19* (1973), 320—326.
3. Loftsgaarden D. O. and Quesenberry C. P. — *A nonparametric estimate of a multivariate density function*, Ann. of Math. Stat. *36* (1965), 1049—1051.
4. Devroye L. P and Wagner T. J. — *The strong uniform consistency of nearest neighbor density estimates*, Ann. of Statistics 5, No. 3 (1977), 536—540.

Received 17. XI.1977

Department of Mathematics
University of Rhode Island
Kingston, Rhode Island 02881
U.S.A.

Analele ştiinţifice ale Universităţii „Al. I. Cuza" Iaşi
Tomul XXIV, f. 2, s. I a, 1978

THEORY OF LOADED COSSERAT ELASTIC CYLINDERS

BY

GH. RUSU

1. Introduction. The Saint-Venant's problem for inhomogeneous and anisotropic cylinders in Cosserat elasticity was solved by D. Ieşan in [3]. In this paper, following [4] we extend these results to the case of loaded cylinders. A Cosserat elastic cylinder acted by body forces, body couples and surface loadings is considered.

2. Statement of the Problem. Throughout this paper, Ω denotes the interior of a right cylinder of length l and with generic cross-section Σ. B is the lateral surface of the cylinder and L the boundary of Σ. A rectangular Cartesian coordinate system, Ox_k, $(k = 1, 2, 3)$ is used. We assume that Ox_3 is parallel to the generators of Ω and x_1Ox_2 plane contains one of the terminal sections. Let $\Sigma^{(0)}$ be the cross-section located at $x_3 = 0$ and $\Sigma^{(l)}$ the cross-section located in the plane $x_3 = l$.

We shall employ the usual summation and differentiation conventions: greek subscripts are understood to run over integers (1,2); latin subscripts are understood to run over integers (1, 2, 3) (unless otherwise is specified); summation over repeated subscripts is implied and subscripts preceded by a comma denote partial differentiation with respect to the corresponding Cartesian coordinate.

In this paper we consider the linear theory of Cosserat elastic solids. The basic equations of this theory are [1]:

— the equilibrium equations

$$t_{ji,j} + f_i = 0, \quad m_{ji,j} + \varepsilon_{ijk} t_{jk} + l_i = 0, \tag{2.1}$$

— the constitutive equations

$$\begin{aligned} t_{ij} &= A_{ijkl}\, e_{kl} + B_{ijkl}\, \varkappa_{kl}, \\ m_{ij} &= B_{klij}\, e_{kl} + C_{ijkl}\, \varkappa_{kl}, \end{aligned} \tag{2.2}$$

— the geometrical equations

$$e_{ij} = u_{j,i} + \varepsilon_{jik}\, \varphi_k, \quad \varkappa_{ij} = \varphi_{j,i}, \tag{2.3}$$

where we have used the notations: u_i — the components of the displacement vector, φ_i-components of the microrotation vector, e_{ij} and $\varkappa_{ij}$-kinematic characteristics of the strain, t_{ij}-components of the stress tensor, m_{ij}-components of the couple-stress tensor, ε_{ijk}-alternating symbol, A_{ijkl}, B_{ijkl}, C_{ijkl}-characteristic coefficients of the material. The elastic coefficients satisfy the symmetry relations