Ann. I. H. Poincaré – PR 38, 6 (2002) 959–971

S0246-0203(02)01132-9/FLA

ROLLING-BALL METHOD FOR ESTIMATING THE BOUNDARY OF THE SUPPORT OF A POINT-PROCESS INTENSITY

MÉTHODE DE ROULEMENT POUR L'ESTIMATION DE LA FRONTIÈRE DU SUPPORT DE L'INTENSITÉ D'UN PROCESSUS PONCTUEL

Peter HALL [a], **Byeong U. PARK** [a,b,1], **Berwin A. TURLACH** [a,c]

[a] *Centre for Mathematics and its Applications, Australian National University, Canberra, ACT 0200, Australia*

[b] *Department of Statistics, Seoul National University, Seoul 151-742, South Korea*

[c] *Department of Mathematics and Statistics, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

Received 5 March 2001, revised 21 January 2002

ABSTRACT. – We suggest a generalisation of the convex-hull method, or 'DEA' approach, for estimating the boundary or frontier of the support of a point cloud. Figuratively, our method involves rolling a ball around the cloud, and using the equilibrium positions of the ball to define an estimator of the envelope of the point cloud. Constructively, we use these ideas to remove lines from a triangulation of the points, and thereby compute a generalised form of a convex hull. The radius of the ball acts as a smoothing parameter, with the convex-hull estimator being obtained by taking the radius to be infinite. Unlike the convex-hull approach, however, our method applies to quite general frontiers, which may be neither convex nor concave. It brings to these contexts the attractive features of the convex hull: simplicity of concept, rotation-invariance, and ready extension to higher dimensions. It admits bias corrections, which we describe and illustrate through implementation.

MSC: primary 62G07; secondary 62H05

Keywords: Bias correction; Confidence band; Curvature; Envelope; Frontier; Productivity analysis; Rotation invariance

RÉSUMÉ. – On suggère une généralisation de la méthode de l'enveloppe convexe pour estimer la frontière du support d'un nuage de points. De manière imagée, notre méthode consiste à faire

[1] The work of the author was supported by KOSEF, the Korean–Australian Cooperative Science Program 1999, and by the Brain Korea 21 Project.

rouler une boule autour du nuage et à utiliser ses positions d'équilibre pour définir un estimateur de l'enveloppe convexe. On peut ainsi construire une enveloppe convexe généralisée, le rayon de la boule jouant le rôle de paramètre de régularisation et l'enveloppe convexe correpondant à un rayon infini. Comparée à celle de l'enveloppe convexe, cette méthode s'applique à des frontières plus générales tout en conservant les même avantages : simplicité conceptuelle, invariance par rotation, extension immédiate aux dimensions supérieures. Elle donne lieu à des corrections de biais que nous décrivons et illustrons dans des implementations.

1. Introduction

The convex-hull estimator of a boundary or frontier is popular in econometrics, where it is a cornerstone of a method known as 'data envelope analysis' or DEA; see for example Charnes et al. [2], Seiford [22], Gijbels et al. [5] and Kneip, Park and Simar [10]. Relative to some of its competitors it has the advantages of being simple in concept, rotation-invariant in definition, and readily extendible to higher dimensions. However, a disadvantage is that it does not apply beyond the case of convex frontiers, and it does not directly involve a smoothing parameter. In this note we suggest a related method which has immediate extension to estimation of general smooth curves, which involves an adjustment for smoothing, and which retains the virtues of convex hull methods. Our approach admits a degree of bias adjustment, particularly when the point cloud under investigation is in two dimensions.

A figurative definition of our method involves rolling a ball around the edge of a point cloud, and taking the estimator of the frontier, $\mathcal{F}$, of the support of the cloud to be 'the trajectory', in some sense, of the ball. The radius of the ball may be interpreted as a smoothing parameter, and can be varied from place to place. The convergence rate of this rolling-ball estimator depends on our definition of the trajectory.

For example, if we ask only that the frontier have a tangent at each point then it is adequate to take the curve traced out by the ball's centre to be our estimator of $\mathcal{F}$. There, a minimax-optimal convergence rate is achieved if the ball's radius gets smaller (at a suitable rate) as the number of points per unit area diverges. However, if the tangent to the frontier varies smoothly, in a way which is differentiable, then that part of the ball that is nearest to the point cloud, in some sense, should be taken as defining the trajectory, and the ball's radius should be kept bounded away from zero as the density of the cloud increases. More explicitly, we suggest defining the trajectory in terms of the polygonal pattern formed by positions of stable equilibrium for the ball.

If the frontier is convex, and if we take the ball's radius to be infinite and define the trajectory as suggested just above, then we obtain exactly the convex-hull estimator, which may therefore be regarded as a special case of a rolling-ball estimator. However, an infinite radius is not appropriate when estimating a general smooth frontier. There, the nearest analogue of the convex-hull estimator is arguably a method based on triangulation of the points, such as the Delaunay triangulation (see e.g. [18, Section 4.3]). We need a method for removing some of the lines in the triangulation, and the rolling-ball approach provides an algorithm for doing just that.

Related work on boundary and frontier estimation includes that of Korostelev and Tsybakov [12], Korostelev, Simar and Tsybakov [13,14], and Mammen and Tsybakov [14] on optimal estimation of sets and frontiers, and that of Hall, Park and Stern [8] on polynomial based methods. Some of the work on frontier estimation assumes Poisson-distributed points, and some assumes a given number, n, of independently-distributed points. There is of course a duality between the two approaches, in which the intensity function of the former is replaced by n multiplied by the common probability density for the latter. First-order asymptotic results are generally the same in both contexts. We shall work in the Poisson setting.

Properties of the convex hull, in the case where the Poisson point process has an unbounded convex domain, are investigated by Nagaev [17]. Results on the number of vertices (and other quantities) of the convex hull of random points are given by Groeneboom [6] and Cabo and Groeneboom [1], who generalise results by Rényi and Sulanke [20,21] and Efron [4].

Section 2 will introduce our methods, including those for bias correction, and Section 3 will describe numerical implementation. Theoretical properties will be summarised in Section 4, with outlines of proofs given in the appendix.

2. Methodology

2.1. Rolling-ball algorithm

Suppose we observe point-process data $\mathcal{X} = \{\xi_1, \xi_2, \ldots\}$ in d-dimensional Euclidean space $\mathbb{R}^d$, and that the intensity function of $\mathcal{X}$ is supported on a compact set bounded by a smooth frontier $\mathcal{F}$ of dimension $d-1$. We wish to estimate $\mathcal{F}$, and suggest the following algorithm.

Let $r > 0$ denote a smoothing parameter, and roll a d-dimensional sphere of radius r around the perimeter of the point cloud. For almost all positions of the sphere (defined with respect to Lebesgue measure) this motion involves pivoting the sphere about a single point in the cloud. However, in some instances, arising with probability 0 if the sphere is placed randomly against the cloud, the sphere touches 2 or more points. When the sphere touches d points it is in a position of stable equilibrium, in the sense that movement into and out of this configuration, in any direction, produces a discontinuity in the derivative of the position of the centre of the sphere. We call these sets of d points 'equilibrium clusters'. If the point cloud is produced randomly in the continuum then, with probability 1, at no time during its rolling motion does the sphere ever touch $d+1$ points simultaneously.

In principle the value of r should be chosen empirically. In the absence of a purely objective procedure, experimentation is suggested. We shall note in Section 2.2, and argue rigorously in Section 4, that the minimax-optimal choice of r is to take r fixed. And we shall observe in Section 4 that taking $r = \infty$ gives the conventional convex-hull estimator.

Each equilibrium cluster defines a $(d-1)$-flat in $\mathbb{R}^d$. Let $\mathcal{G}$ denote that part of this plane bounded by the $\frac{1}{2}d(d-1)$ lines connecting all pairs of the d points; we call it the 'equilibrium face' associated with that particular equilibrium cluster. (Then, $\mathcal{G}$ is a

line, triangle or tetrahedron in the cases $d = 2$, 3 and 4, respectively.) The union of all such faces is a surface, $\widehat{\mathcal{F}}$. Either $\widehat{\mathcal{F}}$, or a smoothed version of it (possibly incorporating a correction for bias), is our approximation to $\mathcal{F}$.

Smoothing in the present context may amount to no more than passing a smooth interpolant through the union of vertices in the sets $\mathcal{G}$. One approach is to use splines, for example in $d = 2$ or 3 dimensions. For $d = 3$ an attractive alternative, making explicit use of the triangulation required to define $\widehat{\mathcal{F}}$, is that suggested by McLain [15].

In asymptotic terms, if the point process is Poisson, if its intensity ν diverges and if the frontier is twice-differentiable, then the optimal size of r is a constant, not converging to 0 as $\nu \to \infty$. This results in $\widehat{\mathcal{F}}$ converging to $\mathcal{F}$ at rate $\mathrm{O}_p(\nu^{-2/(d+1)})$ in a pointwise sense, which is the minimax-optimal rate for frontiers that are differentiable and satisfy a Lipschitz condition of order 1 on the first derivative. (This result may be proved as in Härdle, Park and Tsybakov [9]. See also Korostelev, Simar and Tsybakov [13].)

The case of estimating a production frontier understood as in Grosskopf [7] is a specialisation, and may be formulated as follows. For each i, let X_i denote a $(d-1)$-vector and Y_i be a scalar, and let $\mathcal{X} = \{\xi_1, \xi_2, \ldots\}$ where $\xi_i = (X_i, Y_i)$. It is assumed that the distribution function $F(\cdot \mid x)$ of Y given $X = x$ has an endpoint at $g(x)$, say:

$$F\{g(x) - y \mid x\} \begin{cases} < 1 & \text{for } y > 0, \\ = 1 & \text{for } y \leqslant 0. \end{cases}$$

We wish to estimate the $(d-1)$-dimensional frontier $\mathcal{F}$ defined by $y = g(x)$.

2.2. Alternative rolling-ball methods

Our decision to base the method on equilibrium faces, rather than take a simpler approach, is motivated by a desire to obtain minimax-optimal performance for twice-differentiable frontiers. See Theorem 4.1 below. However, simpler approaches perform optimally when only one derivative is assumed of the frontier.

For example, we may take the estimator of $\mathcal{F}$ to be the locus of the centre of the ball as it rolls around the point cloud. Equivalently, we could centre a ball of radius r at each point in the cloud, and take the union of the spheres as our estimator of $\mathcal{F}$. Alternatively, returning to rolling the ball around $\mathcal{F}$, we could take the locus of any consistently-defined point on the surface of the ball as the estimator of $\mathcal{F}$. For example, we could use the 'lowest' point, provided we can define 'lowness' in terms of some coordinate axis. This is perhaps reasonable in the context of estimating a productivity frontier, where Cartesian coordinates have important physical interpretation. If we take r to be of size $\nu^{-1/d}$, where ν is the intensity of the point process on its support, then the pointwise rate of convergence of the resulting estimator of $\mathcal{F}$ is $\mathrm{O}_p(\nu^{-1/d})$, which is optimal for frontiers that satisfy a Lipschitz condition of order 1.

Heuristically, the reason the equilibrium-cluster method works optimally well for second-order frontiers is that it implicitly estimates the gradient of the frontier at each point. It does this through the equilibrium face, which with high probability is very nearly parallel to the tangent to $\mathcal{F}$ at each point of the latter which is close to the face. No matter what order of bandwidth is used, the methods described in the previous paragraph

fail to achieve minimax-optimal performance for twice-differentiable frontiers, because they do not address the problem of estimating gradient.

The rates of convergence for the equilibrium-cluster method, and for the simpler procedures suggested above, may be shown (using methods similar to those in our Appendix) to equal $O_p(b^2 + s)$ and $O(b + s)$, where terms in b describe the size of bias, terms in s describe the size of the difference between the estimator and its asymptotic mean, and $b = \nu^{-1/(d+1)} r^{1/(d+1)}$ and $s = \nu^{-2/(d+1)} r^{-(d-1)/(d+1)}$. On equating the orders of these quantities, and solving for r, one can see that, in general, asymptotically optimal performance of the equilibrium-cluster method is attained for fixed r, which produces a convergence rate of $O_p(\nu^{-2/(d+1)})$, while optimal performance of the others is achieved with r decreasing to 0 at rate $\nu^{-1/d}$, which gives a convergence rate for these estimators of $\nu^{-1/d}$. The superiority of the equilibrium-cluster approach, in cases where frontier is twice differentiable, is therefore clear.

2.3. Bias correction

First we deal with the case $d = 2$, where relatively simple corrections are possible. They depend on positive constants $w(q)$ and w', defined by $w(q) = E\{W(q)\}$ and $w' = E(W')$, where the random variables $W(q)$ and W' will be introduced in Section 4. (See (4.2) for a definition of $W(q)$.) Table 1 gives approximate values of $w(q)$, computed by simulation, for a range of q's. In the same way we calculated that $w' \approx 1.12$.

The first correction (for $d = 2$) amounts to shifting the frontier estimator $\widehat{\mathcal{F}}$ an amount $(2r)^{-1/3} \widehat{\Lambda}^{-2/3} w(0)$ away from the point cloud, in a direction perpendicular to $\widehat{\mathcal{F}}$, where $\widehat{\Lambda}$ is an estimator of the intensity, Λ, of the point process at the place P on $\mathcal{F}$ where the frontier is being estimated. This adjusts for the effects of bias under the assumption that the frontier is flat at P. It does not correct for the curvature, p, at P, although that could be achieved by changing the shift to $(2r)^{-1/3} \widehat{\Lambda}^{-2/3} w(\hat{p}r)$, where $\hat{p}$ estimates p. These corrections are justified theoretically by Theorem 4.2, of which a more general, d-variate form is the following, which for future reference we call result (R): If $D(p)$ denotes the distance from the frontier estimator to the true frontier, measured perpendicularly to the latter, then as the Poisson intensity Λ diverges, the distribution of $\{r^{d-1}\Lambda^2\}^{1/(d+1)} D(P)$ converges to a distribution that depends on unknowns only through pr, and has expected value $w(pr)$.

When $p > 0$ the frontier is concave upwards, and so represents a 'valley'. Moreover, $pr > 1$ corresponds to the radius of the ball being so large that the ball cannot touch the vertex of the 'valley' while it is rolling. Therefore, we would not use a value of r such that $\hat{p}r > 1$. That is why Table 1 only gives $w(q)$ for $q < 1$. Indeed, $w(1) = E\{W(1)\} = \infty$, even though $W(1) < \infty$ with probability 1; and $W(q) = \infty$ with probability 1 if $q > 1$.

The second correction is designed for the case of a convex-hull approximation to a convex frontier, and amounts to shifting the frontier estimator by $|\hat{p}/2|^{1/3} \widehat{\Lambda}^{-2/3} w'$ in the perpendicular direction, where $\hat{p}$ is an estimator of the curvature of the frontier at the point where the correction is being made. This adjusts for all the bias of the convex-hull estimator, up to terms that are of the same order as $\Lambda^{-2/3}$ multiplied by the larger of the relative errors in the estimators $\widehat{\Lambda}$ and $\hat{p}$. The correction is justified by Theorem 4.3.

Similar corrections may be developed in any number of dimensions, based on generalised versions of Theorems 4.2 and 4.3. However, corrections that involve the curvature, p, now depend on curvatures in several different directions. There are $\frac{1}{2}d(d-1)$ of these, even after the axis system has been rotated so as to be aligned with the tangent plane at the point P of estimation. Thus, explicit corrections for curvature are arguably not attractive in more than two dimensions.

The d-variate analogue of the simple adjustment for tangent when $d = 2$, i.e. of $(2r)^{-1/3}\widehat{\Lambda}^{-2/3}w(0)$, is $r^{-(d-1)/(d+1)}\widehat{\Lambda}^{-2/(d+1)}w_d$, where $\widehat{\Lambda}$ is an estimate of the intensity of the d-variate point process near P, and w_d is an absolute constant, equal to $2^{-1/3}w(0)$ when $d = 1$. Rather than calculate w_d, one may use a Monte Carlo approach, as follows. Conditional on the data, and assuming temporarily that $\mathcal{F}$ is planar at the point P of approximation, generate a homogeneous point process, with intensity $\widehat{\Lambda}$, below a plane passing through the origin O, and compute the rolling-ball estimate (for the given value of r) of $P \equiv O$. Of course, the estimate will be below O. Repeat this procedure a large number of times, and take δ to be the mean distance of the estimates below O. To correct the original estimate $\widehat{\mathcal{F}}$ at a point P, simply shift the estimate a distance δ further away from the point cloud. If it is necessary to vary $(r, \widehat{\Lambda})$ to new values $(r_1, \widehat{\Lambda}_1)$, say, then in view of result (R) given three paragraphs above we should simply multiply δ by $(r/r_1)^{(d-1)/(d+1)}(\widehat{\Lambda}/\widehat{\Lambda}_1)^{2/(d+1)}$.

We may estimate Λ in a locally adaptive way using a histogram-type method, and in the case $d = 2$ we may estimate p by fitting a quadratic locally to the frontier. Details will be given in Section 3.

3. Algorithm

3.1. Implementation of the rolling-ball algorithm

We begin by describing implementation of the rolling-ball algorithm for $d = 2$. First we determine the Delaunay triangulation and the convex hull of the observed points, performing computations in S-PLUS using the Delaunay triangulation package of Turner and Macqueen [23]. Hence, we start with a triangulation defined by all the points and by a polygon identical to the convex hull. We construct the rolling-ball estimator starting from this polygon. At the same time the triangulation is modified by removing and changing some triangles so that in the end only triangles that are 'inside' the rolling ball estimator remain.

Specifically, to construct the rolling-ball estimate for a given value of r we start at a point on the convex hull and move in one direction, say clockwise. Assume that we are at a point P_1 of the polygon. Then we determine the next point, say P_2, on the polygon, and calculate the Euclidean distance between P_1 and P_2. If this distance is greater than $2r$, then it is clear that in the rolling-ball estimate these two points are not connected. Hence, we modify the triangulation by removing the edge between these two points. In removing this edge we are also removing a triangle from the triangulation. The third point of this triangle now becomes part of the polygon that ultimately defines the rolling-ball estimates. After adding this point to the polygon, it becomes the closest point on the polygon to P_1, and we iterate the process just described. Since P_1 is connected to a finite

number of points at the beginning of this process, it is clear that these iterations will either stop if (1) a point with distance less than $2r$ from P_1 is found or (2) all edges connecting P_1 with other points are removed.

If the distance is less than $2r$, we calculate the centre of the circle of radius r on which P_1 and P_2 lie. This circle is uniquely determined by requiring that its centre be on the 'left' of the edge running from P_1 to P_2.

If one or more points (say, the points in $\mathcal{P}$) connected to P_1 or P_2 by the triangulation lie in the interior of the circle, we determine that point in $\mathcal{P}$, say P_3, such that a disc of radius r with its circumference passing through P_1 and P_3 and with its centre on the left side of the edge running from P_1 to P_3, contains neither P_2 nor any point in $\mathcal{P}$ other than P_3. We add P_3 to the polygon since, by construction, the edge connecting P_1 and P_3 belongs to the rolling-ball estimate. The triangulation is updated by removing the edge between P_1 and P_2 (which removes one triangle). Occasionally it may also be necessary to alter other triangles so that no triangle is intersected by the polygon.

If none of the points that are connected to P_1 or P_2 by the triangulation lie in the interior of the circle, then the edge between P_1 and P_2 belongs to the rolling-ball estimates, and we move from P_1 to P_2 and repeat the process just described until we reach the point at which we started the process. The polygon generated at this stage describes the rolling-ball estimate. It is also clear that this process will terminate after a finite number of steps (if it does not stop with an error because r was chosen too small and we would get two disconnected sets).

3.2. Implementing the bias correction

First we describe our Monte Carlo method for computing $w(q) = E\{W(q)\}$ and $w' = E(W')$. (Definitions of $W(q)$ and W' are given in Section 4.) To estimate these quantities we simulate, for each value of q, 500 (say) realisations of a Poisson point process with unit intensity, in the region defined by $y < qx^2$. We use a modified version of the algorithm described by Møller [16], to obtain the first 100, 250 500, 1000 and 5000 (say) points of each realisation. For each realisation we calculate $W(q)$, and the values computed from realisations of the same length (i.e. $100, 250, \ldots, 5000$) are then averaged. There are only minor differences between approximations to $w(q)$ obtained from the realisations for which 5000 points are simulated, and approximations based on shorter sequences; this serves as a check on performance of our methods. Monte Carlo averages over realisations with 5000 points are given in Table 1. Analysis of the data in Table 1 shows that a good approximation to $w(q)$ is given by

$$w(q) \approx 0.75 \log(1-q) - 0.006q + 0.68. \tag{3.1}$$

We use this approximation to implement the bias correction. The value of w' is obtained by similar simulations.

The area of a polygon (convex or otherwise) is easily calculated; see O'Rourke [19, p. 24]. After computing the rolling ball estimate we can determine the area that it circumscribes. An estimate of the intensity $\widehat{\Lambda}$ is now readily obtained by dividing the number of observations by this area. It should be noted that this approach will in general underestimate $\widehat{\Lambda}$, although in our experience that does not pose a problem.

Table 1
Values of $w(q)$. The value of $w(q)$ is given by $w(q) = E\{W(q)\}$, for given values of q, and the random variable $W(q)$ is defined at (4.2)

q	$w(q)$	q	$w(q)$	q	$w(q)$
0.950	−2.0971	−0.250	0.8225	−2.500	1.8899
0.925	−1.5036	−0.300	0.8600	−2.750	1.9360
0.900	−1.0329	−0.400	0.9080	−3.000	1.9841
0.800	−0.4003	−0.500	0.9720	−4.000	2.1936
0.750	−0.1795	−0.600	0.9990	−5.000	2.4864
0.700	−0.0407	−0.700	1.1177	−6.000	2.6669
0.600	0.0821	−0.750	1.1086	−7.500	2.7606
0.500	0.2830	−0.800	1.1140	−8.000	2.8581
0.400	0.3862	−0.900	1.1384	−10.000	3.2194
0.300	0.4808	−1.000	1.1788	−12.500	3.4442
0.250	0.4795	−1.250	1.3134	−15.000	3.7324
0.200	0.5304	−1.500	1.4132	−17.500	3.8875
0.100	0.5814	−1.750	1.5100	−20.000	4.0630
0.000	0.6781	−2.000	1.5060	−25.000	4.4892
−0.100	0.7527	−2.250	1.5929	−30.000	4.8715

If a local estimate of $\widehat{\Lambda}$ is desired then the approach above can be easily adapted. Instead of using the polygon that is given by the rolling ball estimator, we would employ the Delaunay triangulation to determine points that are close to the location where a local estimate of $\widehat{\Lambda}$ is desired, and which lie inside the rolling-ball estimate. We would then use the polygon defined by such points to estimate $\widehat{\Lambda}$.

Finally, to obtain an estimate $\hat{p}$ of the curvature p at a point P we calculate cubic splines $x(t)$ and $y(t)$ such that the curve $(x(t), y(t))$ interpolates the data. The curvature p at a point P is estimated by the curvature of the interpolating curve at P.

Examples of numerical implementation are given in a longer version of this paper, avaliable from the authors.

4. Theoretical properties

We suppose throughout that the point process $\mathcal{X}$ is Poisson with intensity $\Lambda = \nu\lambda$, where λ is a fixed function defined on $\mathbb{R}^d$ and the scalar ν is allowed to diverge to infinity. We assume that:

λ is compactly supported and bounded away from 0 on its support; that the support is a connected set with frontier $\mathcal{F}$; that ball radius, r, is strictly less than the largest

radius such that the ball may roll freely in a neighbourhood of a point P on $\mathcal{F}$ without touching more than one point of $\mathcal{F}$. (4.1)

(For second-order surfaces, such as those assumed in the theorems, the latter assumption will always be valid if r is sufficiently small.) In practice, adaptive smoothing when estimating $\mathcal{F}$ may be achieved by choosing r to be a function of location.

We compute $\widehat{\mathcal{F}}$ as suggested in Section 2.1, using a ball with fixed radius. Given an interior point P of $\mathcal{F}$, let $D(P)$ equal the distance from $\mathcal{F}$ to $\widehat{\mathcal{F}}$, measured perpendicularly to the tangent plane to $\mathcal{F}$ at P. Our next result shows that the minimax-optimal convergence rate, $\mathrm{O}(\nu^{-2/(d+1)})$, is obtained for fixed r.

THEOREM 4.1. – *In addition to assumptions* (4.1), *suppose that in a neighbourhood of P, the first derivatives of the function defining $\mathcal{F}$ exist and satisfy a Lipschitz condition of order* 1. *Then, $D(P) = \mathrm{O}_p(\nu^{-2/(d+1)})$ as $\nu \to \infty$.*

Next we describe the limiting distribution in the case $d = 2$. Let q be any real number. Given a Cartesian coordinate system in $\mathbb{R}^2$ with axes x and y, let $\mathcal{X}_0 = \{(\xi_1, \eta_1), (\xi_2, \eta_2), \ldots\}$ denote a homogeneous Poisson process, with unit intensity, in the region defined by $y < qx^2$. Let $i = I(1)$ be the index that minimises $\alpha_i \equiv \xi_i^2 - \eta_i$ over $i \geqslant 1$. Given $I(k)$ for some $k \geqslant 1$, define

$$\beta_j = \frac{1}{2}\biggl(\xi_{I(k)} + \xi_j - \frac{\eta_{I(k)} - \eta_j}{\xi_{I(k)} - \xi_j}\biggr),$$

and choose $j = I(k+1)$ to minimise $(\xi_{I(k)} - \beta_j)^2$ over

$$j \in \bigl\{j \geqslant 1 \text{ satisfying } (\xi_{I(k)} - \xi_j)\xi_{I(k)} > 0\bigr\}.$$

Let $\hat{k}$ denote the smallest k such that $\xi_{I(k)}\xi_{I(1)} < 0$, and define $I = I(\hat{k} - 1)$, $J = I(\hat{k})$ and

$$W(q) = \xi_I(\eta_I - \eta_J)(\xi_I - \xi_J)^{-1} - \eta_I. \tag{4.2}$$

THEOREM 4.2. – *In addition to assumptions* (4.1), *suppose $d = 2$, that the frontier $\mathcal{F}$ has two continuous derivatives in a neighbourhood of P, that $\lambda(\cdot)$, restricted to its support, is continuous in a neighbourhood of P, and that $\lambda = 1$ at P. Let p denote the curvature at P, with the convention that $p < 0$ or > 0 according as the frontier is concave* (*towards the point cloud*) *or convex at P. Then, $(2r)^{1/3}\nu^{2/3}D(P)$ converges in distribution to $W(pr)$ as $\nu \to \infty$.*

The 'free rolling' condition among assumptions (4.1) guarantees that $W(pr) > 0$ with probability 1. The following definition of $W = W(q)$ is equivalent to the one above, but provides greater geometric insight and is used in the proof of Theorem 4.2. Choose $I(1) = i$ such that the parabola defined by $y = x^2 - \alpha_i$ is as 'high' as possible, subject to containing at least one point of $\mathcal{X}_0$. (That is, move the parabola $y = x^2$ down the y-axis until it first meets a point, which we call $(\xi_{I(1)}, \eta_{I(1)})$.) Given $I(k)$ for some $k \geqslant 1$, choose $I(k+1) = j$ such that (a) $j \neq I(k)$; (b) (ξ_j, η_j) is on the same side of the point $P(k)$, defined to have coordinates $(\xi_{I(k)}, \eta_{I(k)})$, as O; and (c) the parabola with equation

$y = a + (x - b)^2$, where the constants a and b are chosen so that (i) it passes through both $P(k)$ and (ii) the point with coordinates (ξ_j, η_j), is as 'high' as possible. Continue this process until the first time that $P(k)$ is on the opposite side of the y-axis to $P(1)$. Let $\hat{k} \geqslant 2$ be the smallest k for which this is true, and put $I = I(\hat{k} - 1)$ and $J = I(\hat{k})$. Then, the parabola of the form $y = a + (x - b)^2$ that passes through both $P(\hat{k} - 1)$ and $P(\hat{k})$ has $a = \eta_I - (\xi_I - \beta_J)^2$ and $b = \beta_J$, and so the y coordinate of the parabola's vertex is $\eta_I - (\xi_I - \beta_J)^2$. (Thus, the process consists of sliding the parabola downwards, and sideways in the direction of the origin, keeping it touching the latest point $P(k)$, until it first meets a point on the opposite side of the y-axis from $P(1)$.) Let $-W$ equal the point at which the line joining (ξ_I, η_I) and (ξ_J, η_J) cuts the y-axis.

The case of large r is of particular interest, partly because $r = \infty$ and $p < 0$ correspond, at least locally, to a convex-hull approximation to $\mathcal{F}$. First we treat the case $p \geqslant 0$, however. There, the 'free rolling' condition in assumptions (4.1) is important; it requires $pr < 1$ for all sufficiently large r as r increases, and in particular that p should decrease at least as fast as $O(r^{-1})$ if $p > 0$. It may be proved that, if $r = r(\nu) \to \infty$ and $pr \to \ell$, where $\ell \in [0, 1)$, then $\nu^{2/3} D(P) \to 0$ in probability as $\nu \to \infty$. For example, this is the case if $\mathcal{F}$ is flat at the origin, in which setting the result is intuitively clear. (The convergence rate is thus a little faster than the theoretical optimum, this being possible since the curvature is now vanishingly small.)

Our next result addresses the case where $p < 0$ and $r = \infty$. We construct the random variable W' as follows. Redefine $\mathcal{X}_0 = \{(\xi_1, \eta_1), (\xi_2, \eta_2), \ldots\}$ to be a homogeneous Poisson process, with unit intensity, in the region given by $y < -x^2$. Consider the convex hull of $\mathcal{X}_0$ (an estimator of the frontier $y = -x^2$), and let $-W'$ equal the point where the hull crosses the y-axis.

THEOREM 4.3. – *In addition to the assumptions of Theorem* 4.2, *suppose the frontier is concave (towards the point cloud), at least over the region where we are estimating it. Let* $\widehat{\mathcal{F}}$ *be the convex-hull estimator* (*that is, we employ* $r = \infty$), *and let* $p < 0$ *denote the curvature of* $\mathcal{F}$ *at* P. *Then,* $|2/p|^{1/3} \nu^{2/3} D(P) \to W'$ *in distribution as* $\nu \to \infty$.

This result is also valid if, instead of $r = \infty$, $r = r(\nu) \to \infty$ as $\nu \to \infty$, subject to the 'free rolling' condition. Theorem 4.3 is essentially a version in the point-process context of Corollary 1 of Gijbels et al. [5], the main difference being that we give here a constructive definition of the limiting distribution, rather than a formula for its distribution function.

Appendix A

A.1. Proof of Theorem 4.1

Without loss of generality, the point P on $\mathcal{F}$ at which we estimate $\mathcal{F}$ is the origin O, and the tangent plane to $\mathcal{F}$ at P is parallel to the plane of the first $d - 1$ coordinate axes (all but the z-axis, say). The latter assumption is permissible because our estimator $\widehat{\mathcal{F}}$ is invariant under rotations of the data. We shall assume initially that, in a neighbourhood of O, $\mathcal{F}$ is actually planar; and then we shall address the alterations necessary to deal with the more general case.

We first note that $D(P)$ is determined by the equilibrium face which 'cuts' the z-axis. This face is formed by a set of d points which lie on different sides of the first $d-1$ coordinate axes, i.e. have different sign configurations of the first $d-1$ coordinates to each other.

Let $z_\nu = \nu^{-2/(d+1)}$. Given a constant $C > 0$, consider a d-dimensional rectangle

$$\mathcal{R} = \bigl\{(x_1, \ldots, x_{d-1}, z)\colon -\bigl(rCz_\nu/(d-1)\bigr)^{1/2} \leqslant x_i \leqslant \bigl(rCz_\nu/(d-1)\bigr)^{1/2}, \\ -Cz_\nu/2 \leqslant z \leqslant 0\bigr\}.$$

Partition $\mathcal{R}$ into 2^{d-1} sub-rectangles according to the 2^{d-1} sign configurations of the first $d-1$ coordinates. Denote them by $\mathcal{R}_i$, $i = 1, \ldots, 2^{d-1}$. For example, in the case $d = 3$ one of the sub-rectangles is given by

$$\bigl\{(x_1, x_2, z)\colon 0 \leqslant x_1, x_2 \leqslant (rCz_\nu/2)^{1/2},\ -Cz_\nu/2 \leqslant z \leqslant 0\bigr\}.$$

Suppose there exists at least 1 point in each of the partitions $\mathcal{R}_i$. Then the protrusion below the plane $z = 0$ of a d-dimensional sphere of radius r in the position of the equilibrium which defines $D(P)$, never exceeds Cz_ν. The maximum protrusion occurs when the sphere touches a particular set of d points among the 'lower' 2^{d-1} vertices of $\mathcal{R}$. This means that $D(P) \leqslant Cz_\nu$.

Let $C_1 > 0$ be a lower bound to λ on its support $\mathcal{S}$. It follows then that

$$\begin{aligned} P\bigl\{D(P) > Cz_\nu\bigr\} &\leqslant P\bigl\{\text{No points in } \mathcal{R}_i \text{ for some } 1 \leqslant i \leqslant 2^{d-1}\bigr\} \\ &\leqslant \sum_{i=1}^{2^{d-1}} \exp\biggl\{-\int_{\mathcal{R}_i} \nu\lambda(\xi)\,d\xi\biggr\} \\ &\leqslant 2^{d-1}\exp\bigl[-C_1\nu\bigl\{rCz_\nu/(d-1)\bigr\}^{(d-1)/2}(Cz_\nu/2)\bigr] \\ &= 2^{d-1}\exp\bigl[-C_1\{r/(d-1)\}^{(d-1)/2}C^{(d+1)/2}/2\bigr], \end{aligned}$$

which tends to zero as $C \to \infty$.

If a portion z, measured radially, of a d-dimensional sphere $\mathcal{T}$ of radius r protrudes below a plane, then the radius of the $(d-1)$-dimensional sphere formed by the intersection of the plane with $\mathcal{T}$, equals $\mathrm{O}(z^{1/2})$ as $z \downarrow 0$. Therefore, if $\mathcal{F}$ is not planar in a neighbourhood of 0, the fact that the tangent plane satisfies a Lipschitz condition of order 1 as it is moved around $\mathcal{F}$ implies that $D(P)$ differs by no more than $\mathrm{O}\{(z_\nu^{1/2})^2\} = \mathrm{O}(z_\nu)$ from its position in the planar case. Thus, the result continues to hold.

A.2. Proof of Theorem 4.2

We may suppose that $\mathcal{F}$ passes through $(0, 0)$, that its tangent at that point is the line $y = 0$, and that the point cloud is below $\mathcal{F}$. We assume too that the point process has intensity identically equal to ν; the case where the intensity equals $\nu\lambda(\cdot)$, and $\lambda(x, y) \to 1$ as $(x, y) \to (0, 0)$, may be treated similarly.

Suppose the ball (here a disc) is centred at $(x_1, y_1) \equiv (c_1\theta + \mathrm{O}(\theta^2), r + c_2\theta^2 + \mathrm{O}(\theta^3))$, where $\theta > 0$ is small and $-\infty < c_1, c_2 < \infty$. (We do not include terms of size θ in the expansion of y_1, since if the ball has a protrusion of width $\mathrm{O}(\theta)$ below $\mathcal{F}$ then the

depth of that protrusion will be $O(\theta^2)$, not just $O(\theta)$.) The circumference of the ball has equation $(x-x_1)^2+(y-y_1)^2=r^2$, which implies that $y/r=\frac{1}{2}\{(x/r)+d_1\theta\}^2+d_2\theta^2+O(|x|^3+\theta^3)$ as $\theta+|x|\to 0$, for constants d_1, d_2 determined by c_1, c_2. Re-parametrising to $x=hru$, $y=\frac{1}{2}h^2rv$ and $\theta=ht$, where $h=\{2/(r^2\nu)\}^{1/3}$, we obtain

$$v=a+(u-b)^2+O(h) \tag{A.3}$$

as $h\to 0$, where a, b depend on c_1, c_2, t. (The order of the remainder term is valid provided $|u|=O(1)$.)

If the curvature, or second derivative, of $\mathcal{F}$ at $(0,0)$ equals p then the locus of points (x, y) on $\mathcal{F}$ has equation $y=\frac{1}{2}px^2+O(|x|^3)$ as $x\to 0$. Reparametrising as before, the equation becomes

$$v=pru^2+O(h) \tag{A.4}$$

as $h\to 0$, assuming that $|u|=O(1)$.

The intensity of the Poisson process in (u, v)-space equals 1. Therefore, in the limit as $h\to 0$ (or equivalently, as $\nu\to\infty$), the problem of rolling a ball across the top of a point cloud (emanating from a Poisson process with intensity ν, below the frontier $\mathcal{F}$) near the origin O, until it just touches two points, converges to one of 'sliding' a solid parabola, whose perimeter has Eq. (A.3), across the cloud so that it just touches two points of another cloud (this time coming from a Poisson process with unit intensity, and distributed below the frontier defined by (A.4)) near the origin.

The latter point process, and parabola-sliding algorithm, is exactly the one used to define the distance $W(q)$, with $q=rp$, of O from the point on the parabola immediately below O. See the second definition of $W(q)$ in Section 4. Hence, after re-parametrisation to the (u, v)-plane, the distance below the origin of the nearest equilibrium face (here, a line) converges in distribution to $W(pr)$. Equivalently, returning to the scale of the original coordinate system, $D(O)/(\frac{1}{2}h^2r)$ converges in distribution to $W(pr)$ as $\nu\to\infty$. This is equivalent to Theorem 4.2.

REFERENCES

[1] A.J. Cabo, P. Groeneboom, Limit theorems for functionals of convex hulls, Probab. Theory Related Fields 100 (1994) 31–55.
[2] A. Charnes, W.W. Cooper, A.Y. Lewin, L.M. Seiford, Data Envelope Analysis: Theory, Methodology and Applications, Kluwer, Boston, 1995.
[3] L. Christensen, R. Greene, Economics of scale in US electric power generation, J. Polit. Economy 84 (1976) 653–667.
[4] B. Efron, The convex hull of a random set of points, Biometrika 52 (1965) 331–343.
[5] I. Gijbels, E. Mammen, B.U. Park, L. Simar, On estimation of monotone and concave frontier functions, J. Amer. Statist. Assoc. 94 (1999) 220–228.
[6] P. Groeneboom, Limit theorems for convex hulls, Probab. Theory Related Fields 79 (1988) 327–368.
[7] S. Grosskopf, Statistical inference and nonparametric efficiency: a selective survey, J. Productivity Anal. 7 (1996) 161–176.

[8] P. Hall, B.U. Park, S. Stern, On polynomial estimators of frontiers and boundaries, J. Multivariate Anal. 66 (1998) 71–98.
[9] W. Härdle, B.U. Park, A.B. Tsybakov, Estimation of non-sharp support boundaries, J. Multivariate Anal. 55 (1995) 205–218.
[10] A. Kneip, B.U. Park, L. Simar, A note on the convergence of nonparametric DEA estimators for production efficiency scores, Econometric Theory 14 (1998) 783–793.
[11] A.P. Korostelev, A.B. Tsybakov, Minimax Theory of Image Reconstruction, in: Lecture Notes in Statistics, Vol. 82, Springer-Verlag, Berlin, 1993.
[12] A.P. Korostelev, L. Simar, A.B. Tsybakov, Efficient estimation of monotone boundaries, Ann. Statist. 23 (1995) 476–489.
[13] A.P. Korostelev, L. Simar, A.B. Tsybakov, On estimation of monotone and convex boundaries, Pub. Inst. Statist. Univ. Paris 49 (1995) 3–18.
[14] E. Mammen, A.B. Tsybakov, Asymptotical minimax recovery of sets with smooth boundaries, Ann. Statist. 23 (1995) 502–524.
[15] D.H. McLain, Two dimensional interpolation from random data, Comput. J. 19 (1976) 178–181.
[16] J. Møller, Lectures on Random Voronoi Tessellations, in: Lecture Notes in Statistics, Vol. 87, Springer-Verlag, New York, 1994.
[17] A.V. Nagaev, Some properties of convex hulls generated by homogeneous Poisson point processes in an unbounded convex domain, Ann. Inst. Statist. Math. 47 (1995) 21–29.
[18] B.D. Ripley, Spatial Statistics, Wiley, New York, 1981.
[19] J. O'Rourke, Computational Geometry in C, Cambridge University Press, Cambridge, 1994.
[20] A. Rényi, R. Sulanke, On the convex hull of n randomly chosen points, Z. Wahrscheinlichkeitstheorie Verw. Geb. 2 (1963) 75–84.
[21] A. Rényi, R. Sulanke, On the convex hull of n randomly chosen points II, Z. Wahrscheinlichkeitstheorie Verw. Geb. 3 (1964) 138–147.
[22] L.M. Seiford, Data envelopment analysis: the evolution of the state-of-the-art, 1978–1995, J. Productivity Anal. 7 (1996) 99–137.
[23] R. Turner, D. Macqueen, S function Deldir to compute the Dirichlet (Voronoi) tesselation and Delaunay triangulation of a planar set of data points, Available from Statlib, 1996.