On estimation of the intensity function of a point process

M.N.M. van Lieshout[1]

[1] CWI and Eindhoven University of Technology,
Science Park 123, 1098 XG Amsterdam, The Netherlands

Abstract. Estimation of the intensity function of spatial point processes is a fundamental problem. In this paper, we interpret the Delaunay tessellation field estimator recently introduced by Schaap and Van de Weygaert as an adaptive kernel estimator and give explicit expressions for the mean and variance. Special attention is paid to Poisson processes.

Keywords. Delaunay tessellation field estimator, generalised weight function estimator, intensity function, kernel estimator, mass preservation, Poisson process, second order product density.

1 Introduction

Let Φ be a simple point process on $\mathbb{R}^d$ whose moment measure M exists and is locally finite. Make the further assumption that M is absolutely continuous with respect to Lebesgue measure, so that

$$M(A) = \int_A \lambda(x)\, dx$$

for some measurable function $\lambda(x) \geq 0$ for every bounded Borel set A. The goal is to estimate $\lambda(\cdot)$ based on a realisation $\Phi \cap A$ of Φ in some open, convex and bounded Borel set $A \neq \emptyset$.

The classic approach is to use the Berman–Diggle (1989) estimator

$$\widehat{\lambda_{BD}(x_0)} := \frac{N(b(x_0,h) \cap A)}{|b(x_0,h) \cap A|}, \tag{1}$$

for $x_0 \in A$. Here the notation $N(B)$ is used for the cardinality of $\Phi \cap B$ for bounded Borel sets $B \subset \mathbb{R}^d$, and $|B|$ for its Lebesgue mass. The bandwidth $h > 0$ controls the amount of smoothing. Note that as A is open, one never divides by zero. By construction, the expectation of (1) is $M(b(x_0,h) \cap A)/|b(x_0,h) \cap A|$.

Although (1) is a natural estimator, it does not necessarily preserve the total mass in W, nor is it based on a (generalised) weight function. In this paper we consider alternative estimators that do possess these two properties.

2 Mass preserving kernel estimation

Note that (1) can be regarded as a kernel estimator with kernel $1\{||x - x_0|| < h\}/|b(x_0,h) \cap A|$. The edge correction term $|b(x_0,h) \cap A|$ is 'global' as it does not depend on x. Using a 'local' edge correction instead suggests the estimator

$$\widehat{\lambda_K(x_0)} := \sum_{x \in \Phi \cap A} \frac{1\{||x - x_0|| < h\}}{|b(x,h) \cap A|} \tag{2}$$

based on the kernel $k_h(x_0 \mid x) = 1\{||x - x_0|| < h\}/|b(x,h) \cap A|$ for $x_0 \in A$. Note that (2) is well-defined and coincides with (1) for x_0 such that $b(x_0, 2h) \subseteq A$. In contrast to (1), however,

(2) is based on a proper weight function as

$$\int_A k_h(x_0 \mid x)\, dx_0 = \int_A \frac{1\{||x - x_0|| < h\}}{|b(x,h) \cap A|}\, dx_0 \equiv 1$$

for all $x \in A$. Consequently, (2) is mass preserving, that is, $\int_A \widehat{\lambda_K(x_0)}\, dx_0 = N(A)$ almost surely.

The first two moments of (2) are given by

$$\mathbb{E}\left[\widehat{\lambda_K(x_0)}\right] = \int_{A \cap b(x_0,h)} \frac{\lambda(x)}{|b(x,h) \cap A|}\, dx;$$

$$\mathbb{E}\left[\widehat{\lambda_K(x_0)}^2\right] = \int_{(b(x_0,h) \cap A)^2} \frac{\rho^{(2)}(x,y)}{|b(x,h) \cap A|\, |b(y,h) \cap A|}\, dx\, dy + \int_{b(x_0,h) \cap A} \frac{\lambda(x)}{|b(x,h) \cap A|^2}\, dx,$$

provided the second order factorial moment measure of Φ exists as a locally finite measure that is absolutely continuous with respect to the product Lebesgue measure with Radon–Nikodym derivative $\rho^{(2)}$.

3 Local versus global edge correction

Neither (1) nor (2) is universally better in terms of integrated mean squared error than its competitor. To see this, first consider a homogeneous Poisson process Φ with intensity $\lambda > 0$. Then, the integrated variance of both (1) and (2) is equal to $\lambda \int_A |b(x,h) \cap A|^{-1}\, dx$. The bias of the Berman–Diggle estimator is zero, whereas (2) is biased unless $\int_{b(x_0,h) \cap A} |b(x,h) \cap A|^{-1}\, dx = 1$. So, in general, (1) will be preferred.

Next, let Φ be a Poisson process on A with intensity function $\lambda(x) = \lambda\, |b(x,h) \cap A|$, for some $\lambda > 0$. Then, (2) is unbiased with integrated variance $\lambda\, |A|$. Write

$$m(x_0) := \int_{b(x_0,h) \cap A} \frac{|b(x,h) \cap A|}{|b(x_0,h) \cap A|^2}\, dx.$$

Then, $\mathbb{E}\, \widehat{\lambda_{BD}(x_0)}$ can be expressed as $\lambda(x_0)\, m(x_0)$, so its integrated squared bias is zero if and only if $m(x_0) = 1$ for almost all $x_0 \in A$. The integrated variance of (1) is $\lambda \int_A m(x_0)\, dx_0$ which reduces to $\lambda |A|$ if $m(x_0) = 1$ for almost all $x_0 \in A$ so that the estimators are indistinguishable. Otherwise the mass preserving kernel estimator should be preferred since $\int_A m(x_0)\, dx_0 \geq |A|$.

4 Delaunay tessellation field estimator

Suppose that realisations of the point process Φ are almost surely in general quadratic position, that is, no $d+2$ points are located on the boundary of a sphere and no $k+1$ points lie in a $k-1$ dimensional affine subspace for $k = 2, \dots d$. Then the Delaunay tessellation of Φ is well defined. The union of Delaunay cells containing a point $x_i \in \Phi$ is the contiguous Voronoi cell of x_i in Φ and will be denoted by $W(x_i \mid \Phi)$. For further details, see e.g. Møller (1994) or Okabe et al. (2000).

Note that the tessellation cells described above can be seen as adaptive neighbourhoods of a point of Φ. In contrast to the balls of fixed radius h used before, the size of the cells depend on the point process: In densely populated regions, the cells will be small, whereas they tend to be larger in regions of low intensity. Based on this idea, Schaap and Van de Weygaert (2000,

2007) introduced the Delaunay tessellation field estimator (DTFE) as follows. For $x \in \Phi \cap A$, set $\widehat{\lambda_D(x)} := (d+1)/|W(x \mid \Phi \cap A)|$, and for any $x_0 \in A$ in the interior of some Delaunay cell, define

$$\widehat{\lambda_D(x_0)} := \frac{1}{d+1} \sum_{x \in \Phi \cap D(x_0 \mid \Phi \cap A)} \widehat{\lambda_D(x)} \tag{3}$$

by averaging over the vertices of the Delaunay cell $D(x_0 \mid \Phi \cap A)$ that contains x_0.

The DTFE preserves total mass and is an adaptive kernel estimator. To see this, write $\mathcal{D}(\varphi \cap A)$ for the family of Delaunay cells of $\varphi \cap A$, and set

$$g(x_0 \mid x, \varphi) := \frac{\sum_{D_j \in \mathcal{D}(\varphi \cap A)} 1\{x_0 \in D_j^\circ; x \in D_j\}}{|W(x \mid \varphi \cap A)|},$$

for $x_0 \in A \setminus \varphi$, $x \in \varphi$, and $g(x \mid x, \varphi) := (d+1)/|W(x \mid \varphi \cap A)|$ if $x \in \varphi \cap A$. Then $\widehat{\lambda_D(x_0)} = \sum_{x \in \Phi \cap A} g(x_0 \mid x, \Phi)$, and, as $\int_A g(x_0 \mid x, \varphi)\, dx_0 = 1$, mass preservation follows. Note that there is no need for a subjective choice of bandwidth, though at some computational cost.

The first two moments of (3) are given by

$$\mathbb{E}\left[\widehat{\lambda_D(x_0)}\right] = \int_A \mathbb{E}_x\left[g(x_0 \mid x, \Phi)\right] \lambda(x)\, dx$$

$$\mathbb{E}\left[\widehat{\lambda_D(x_0)}^2\right] = \int_{A^2} \mathbb{E}^{(2)}_{x,y}\left[g(x_0 \mid x, \Phi)\, g(x_0 \mid y, \Phi)\right] \rho^{(2)}(x,y)\, dx\, dy + \int_A \mathbb{E}_x\left[g^2(x_0 \mid x, \Phi)\right] \lambda(x)\, dx$$

provided the second order factorial moment measure of Φ exists as a locally finite measure that is absolutely continuous with respect to the product Lebesgue measure with Radon–Nikodym derivative $\rho^{(2)}$. Here $\mathbb{E}_x$ ($\mathbb{E}^{(2)}_{x,y}$) denotes expectation with respect to the Palm distribution of Φ at x (the two-fold Palm distribution at x, y). If Φ is a Poisson process with intensity function $\lambda(\cdot)$, $P^!_x = P$ and $\rho^{(2)}(x,y) = \lambda(x)\,\lambda(y)$. The result should be compared to its kernel estimation counterpart.

Finally, assume that Φ is a stationary Poisson process on $\mathbb{R}^d$. Then, (3) is unbiased with variance $c_d \lambda^2$. The constant c_d depends on the dimension. For example if $d = 1$, $c_1 = 2\,(2 - \pi^2/6) \approx 0.7$. Since the Berman–Diggle estimator is unbiased with variance $\lambda\, \omega_d^{-1}\, h^{-d}$, where ω_d is the volume of the unit ball, it is more efficient than (3) whenever $\mathbb{E}N(b(0,h)) > 1/c_d$. Hence on the line, (1) is the better choice if $\mathbb{E}N((-h,h)) = 2\lambda h > 1.4$. For comparison, the computation of (3) in this case requires four points. Simulations by the author indicate that DTFE would be preferred for strongly oscillating intensity functions in contexts where peak preservation is important; for mildly fluctuating intensity functions, kernel estimation seems more efficient.

References

Berman, M. and Diggle, P.J. (1989): Estimating Weighted Integrals of the Second-Order Intensity of a Spatial Point Process. *Journal of the Royal Statistical Society Series B*, *51*, 81–92.

Møller, J (1994): *Lectures on Random Voronoi Tessellations.* Lecture Notes in Statistics 87. Springer-Verlag.

Okabe, A. and Boots, B. and Sugihara, K. and Chiu, S.N. (2000): *Spatial Tessellations. Concepts and Applications of Voronoi Diagrams.* Second Edition. Wiley.

Schaap, W.E. (2007): *DTFE. The Delaunay Tessellation Field Estimator.* Ph.D. Thesis, University of Groningen.

Schaap, W.E. and Weygaert, R. van de (2000): Letter to the Editor. Continuous Fields and Discrete Samples: Reconstruction through Delaunay Tessellations. *Astronomy and Astrophysics*, *363*, L29–L32.