Statistics and Probability Letters 133 (2018) 59–64

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

A fast spectral quasi-likelihood approach for spatial point processes

C. Deng [a], R.P. Waagepetersen [b], M. Wang [c], Y. Guan [c,*]

[a] Program in Applied Mathematics, Yale University, New Haven, CT, 06511, USA
[b] Department of Mathematical Sciences, Aalborg University, Fredrik Bajersvej 7G, DK-9220 Aalborg, Denmark
[c] Department of Management Science, University of Miami, Coral Gables, FL, 33124, USA

ARTICLE INFO

Article history:
Received 24 September 2016
Received in revised form 25 September 2017
Accepted 29 September 2017
Available online 24 October 2017

Keywords:
Estimating function
Quasi-Likelihood
Spatial point process
Spectral approach

ABSTRACT

In applications of spatial point processes, it is often of interest to fit a parametric model for the intensity function. For this purpose Guan et al. (2015) recently introduced a quasi-likelihood type estimating function that is optimal in a certain class of first-order estimating functions. However, depending on the choice of certain tuning parameters, the implementation suggested in Guan et al. (2015) can be very demanding both in terms of computing time and memory requirements. Using a novel spectral representation, we construct in this paper an implementation that is computationally much more efficient than the one proposed in Guan et al. (2015).

1. Introduction

Consider a spatial point process observed on a window $W \subset \mathbb{R}^2$. Let $N(B)$ denote the number of events in a Borel set $B \subset \mathbb{R}^2$. Let $\Delta\mathbf{u}$ be a small region containing $\mathbf{u}$ for $\mathbf{u} \in \mathbb{R}^2$. Assuming that the following two limits exist, the first- and second-order intensity functions of the process are defined as

$$\lambda(\mathbf{u}) = \lim_{|\Delta\mathbf{u}|\to 0} \frac{\mathrm{E}\{N(\Delta\mathbf{u})\}}{|\Delta\mathbf{u}|}, \quad \lambda_2(\mathbf{u}_1, \mathbf{u}_2) = \lim_{|\Delta\mathbf{u}_1|,|\Delta\mathbf{u}_2|\to 0} \frac{\mathrm{E}\{N(\Delta\mathbf{u}_1)N(\Delta\mathbf{u}_2)\}}{|\Delta\mathbf{u}_1||\Delta\mathbf{u}_2|}, \tag{1}$$

where $\lambda(\cdot)$ is often called the intensity function. A spatial point process is said to be second-order intensity reweighted stationary (Baddeley et al., 2000) if $\lambda_2(\mathbf{u}_1, \mathbf{u}_2) = \lambda(\mathbf{u}_1)\lambda(\mathbf{u}_2)g(\mathbf{u}_1 - \mathbf{u}_2)$, where $g(\cdot)$ is the pair correlation function (Møller and Waagepetersen, 2004, p. 31). We assume that $\lambda(\mathbf{u})$ is given by a parametric model $\lambda(\mathbf{u}; \boldsymbol{\beta})$ depending on a $p \times 1$ parameter vector $\boldsymbol{\beta}$, where $\lambda(\mathbf{u}; \boldsymbol{\beta})$ is positive and continuously differentiable with respect to $\boldsymbol{\beta}$.

Guan et al. (2015) recently proposed to estimate $\boldsymbol{\beta}$ by solving an estimating equation of the form

$$\mathbf{U}(\boldsymbol{\beta}) = \sum_{\mathbf{u}\in W\cap N} \boldsymbol{\phi}(\mathbf{u}) - \int_W \boldsymbol{\phi}(\mathbf{s})\lambda(\mathbf{s}; \boldsymbol{\beta})d\mathbf{s} = \mathbf{0}, \tag{2}$$

* Corresponding author.
E-mail addresses: chong.deng@yale.edu (C. Deng), rw@math.aau.dk (R.P. Waagepetersen), mwang@bus.miami.edu (M. Wang), yguan@bus.miami.edu (Y. Guan).

https://doi.org/10.1016/j.spl.2017.09.016
0167-7152/

where $\boldsymbol{\phi}(\cdot)$ is a weight function. They showed that the optimal choice of the weight function is the solution of the Fredholm integral equation:

$$\boldsymbol{\phi}(\mathbf{u}) = \frac{\boldsymbol{\lambda}^{(1)}(\mathbf{u};\boldsymbol{\beta})}{\lambda(\mathbf{u};\boldsymbol{\beta})} - \mathbf{T}\boldsymbol{\phi}(\mathbf{u}), \tag{3}$$

where $\boldsymbol{\lambda}^{(1)} = d\lambda/d\boldsymbol{\beta}$ and $\mathbf{T}$ is an integral operator given by

$$\mathbf{T}\boldsymbol{\phi}(\mathbf{u}) = \int_W R(\mathbf{u}-\mathbf{v})\lambda(\mathbf{v};\boldsymbol{\beta})\phi(\mathbf{v})d\mathbf{v} \tag{4}$$

with a kernel function $R(\cdot) = g(\cdot) - 1$. The resulting estimating function can be interpreted as a quasi-likelihood score. It is optimal since it achieves the maximal Godambe information and hence the minimal asymptotic variance among all estimating functions of the form (2).

To solve (3), Guan et al. (2015) employed the Nyström method by solving a system of linear equations,

$$\boldsymbol{\phi}(\mathbf{u}_j) + \sum_{k=1}^{m} R(\mathbf{u}_j - \mathbf{u}_k)\lambda(\mathbf{u}_k;\boldsymbol{\beta})\boldsymbol{\phi}(\mathbf{u}_k)w_k = \frac{\boldsymbol{\lambda}^{(1)}(\mathbf{u}_j;\boldsymbol{\beta})}{\lambda(\mathbf{u}_j;\boldsymbol{\beta})}, \quad j = 1,\ldots,m, \tag{5}$$

where $\mathbf{u}_j$ and $w_j, j = 1,\ldots,m$, are quadrature points and quadrature weights used to approximate the integral in (4). The solution to (5) is then used in a quadrature approximation of (4) to estimate $\mathbf{T}\boldsymbol{\phi}(\mathbf{u})$ and via (3) the weight function $\boldsymbol{\phi}(\mathbf{u})$ at an arbitrary location $\mathbf{u}$.

The Nyström method becomes more accurate with an increasing number of quadrature points m. However, the method becomes computationally challenging when m is large. In particular it is required to store and invert an $m \times m$ covariance matrix with entries given in terms of $R(\mathbf{u}_j - \mathbf{u}_k)$. Guan et al. (2015) proposed a tapering scheme where the jk'th entry is replaced by zero if the inter-point distance between $\mathbf{u}_j$ and $\mathbf{u}_k$ is beyond a certain threshold. However, the amount of storage can still be enormous and the computational time needed to solve (5) may increase at a rate up to order m^3 depending on the actual value of the threshold used. This is problematic since a large m may be needed in practice in order to adequately capture the varying intensity and/or dependence in the spatial point process, or to investigate the sensitivity of the parameter estimates to a range of m values.

In the next section, we introduce a novel approach that overcomes the challenges faced by the standard Nyström method. We first develop a novel spectral representation of the kernel function $R(\cdot)$ used in (4) based on which we express $R(\cdot)$ as a so-called degenerate kernel in the integral equation literature (Jerri, 1985, p. 123). We next convert the Fredholm integral equation (3) to a new Fredholm equation in the spectral domain. The Nyström method is then applied to solve the spectral Fredholm equation. Our theoretical results in Section 2.3 regarding computational complexity and memory requirements as well as the simulation studies show that the new spectral implementation can greatly outperform the existing one when the grid size m is large. Variance estimation of the resulting estimator can be done extremely fast using the solution to the spectral Fredholm equation. This is another advantage of our approach over that in Guan et al. (2015), where a double integral involving $R(\cdot)$ has to be numerically evaluated in order to estimate the variance.

2. The new spectral approach

2.1. A Fredholm integral equation in the spectral domain

Following Jalilian et al. (2013) we assume that $R(\cdot)$ is a positive definite function. This indeed holds e.g. for wide classes of Cox and cluster point processes. More specifically we assume that $R(\cdot)$ belongs to a parametric family of covariance functions characterized by a parameter vector $\boldsymbol{\theta}$. For example, for the flexible class of Matérn covariance models (Jalilian et al., 2013),

$$R(\mathbf{h};\boldsymbol{\theta}) = \omega_0^2 \frac{\left(\sqrt{2\nu}\|\mathbf{h}\|/\mathbf{2}\sigma\right)^{\nu} K_\nu\left(\sqrt{2\nu}\|\mathbf{h}\|/\mathbf{2}\sigma\right)}{2^{\nu-1}\Gamma(\nu)}, \text{ for } \mathbf{h} \in \mathbb{R}^2, \tag{6}$$

where $\Gamma(\cdot)$ is the gamma function, $K_\nu(\cdot)$ is the modified Bessel function of the second kind, $\boldsymbol{\theta} = (\kappa,\nu,\sigma)^T$, $\omega_0^2 = 1/(16\pi\sigma^2\nu\kappa)$, and κ, σ and ν are non-negative parameters.

By Bochner's theorem (Bochner, 1955),

$$R(\mathbf{h};\boldsymbol{\theta}) = \int_{\mathbb{R}^2} f(\boldsymbol{\omega};\boldsymbol{\theta})\exp(-2\pi i\boldsymbol{\omega}'\mathbf{h})d\boldsymbol{\omega}, \tag{7}$$

where $f(\cdot) > 0$ is the spectral density of $R(\cdot)$ and i is the imaginary unit. For Matérn covariance models,

$$f(\boldsymbol{\omega};\boldsymbol{\theta}) = \frac{4\pi\omega_0^2\Gamma(\nu+1)(2\nu)^\nu}{\Gamma(\nu)(2\sigma)^{2\nu}}\left(\frac{\nu}{2\sigma^2} + 4\pi^2\|\boldsymbol{\omega}\|^2\right)^{-\nu-1}.$$

By (7),

$$R(\mathbf{u}-\mathbf{v};\boldsymbol{\theta}) = \int_{\mathbb{R}^2} f(\boldsymbol{\omega};\boldsymbol{\theta})\exp(2\pi i\boldsymbol{\omega}'\mathbf{u})\exp(-2\pi i\boldsymbol{\omega}'\mathbf{v})d\boldsymbol{\omega}. \tag{8}$$