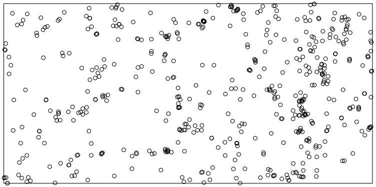
Estimation of the pair correlation function of a spatial point process

Rasmus Waagepetersen
Department of Mathematics
Aalborg University
Denmark

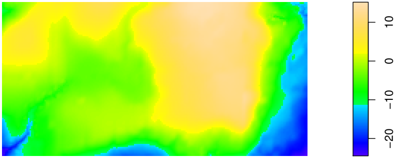
June 21, 2017

- joint work with Abdollah Jalilian and Yongtao Guan

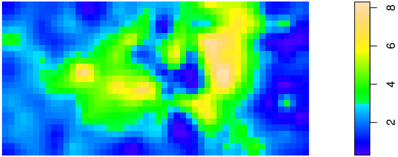
Data example: *Acalypha*



- *observation window W* = 1000 m × 500 m
- seed dispersal ⇒ *clustering*
- environment ⇒ *inhomogeneity*



Elevation



Potassium content in soil

Objective: quantify dependence on environmental variables and clustering

Spatial point process

A locally finite random subset $\mathbf{X}$ of $\mathbb{R}^2$:

$\#(\mathbf{X} \cap A)$ finite for all bounded subsets $A \subset \mathbb{R}^2$

We let $N(A) = \#(\mathbf{X} \cap A)$ for bounded $A \subset \mathbb{R}^2$.

Moments of a spatial point process

Fundamental characteristics of point process: mean and covariance of counts $N(A) = \#(\mathbf{X} \cap A)$.

Intensity measure μ:

$$\mu(A) = \mathbb{E}N(A), \quad A \subseteq \mathbb{R}^2$$

In practice often given in terms of *intensity function*

$$\mu(A) = \int_A \rho(\mathbf{u}) \mathrm{d}\mathbf{u}$$

We further assume that covariances given in terms of intensity function and pair correlation function:

$$\mathbb{C}\mathrm{ov}[N(A), N(B)] = \int_{A\cap B} \rho(\mathbf{u})\mathrm{d}\mathbf{u} + \int_A \int_B \rho(\mathbf{u})\rho(\mathbf{v})(g(\mathbf{u},\mathbf{v}) - 1)\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{v}$$

$$= \text{ Poisson covariance } + \text{additional/less covariance due to interaction}$$

Further interpretation of g: let $\rho(\mathbf{u}|\mathbf{v})$ denote intensity of $\mathbf{X}$ given $\mathbf{v} \in \mathbf{X}$ ('Palm' intensity). Then

$$g(\mathbf{u},\mathbf{v}) = \frac{\rho(\mathbf{u}|\mathbf{v})}{\rho(\mathbf{u})}$$

- how much is intensity of **u** changed by presence of point at **v**.

Campbell formulae

By definition of intensity function and product density and the standard proof we obtain the useful Campbell formulae:

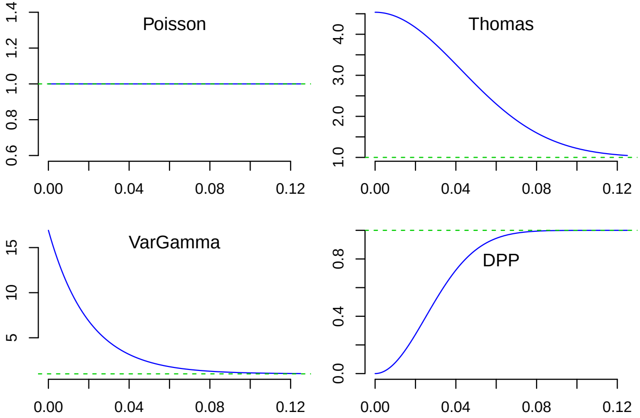
$$\mathbb{E} \sum_{\mathbf{u} \in \mathbf{X}} h(\mathbf{u}) = \int h(\mathbf{u}) \rho(\mathbf{u}) \mathrm{d}\mathbf{u}$$

$$\mathbb{E} \sum_{\mathbf{u}, \mathbf{v} \in \mathbf{X}}^{\neq} h(\mathbf{u}, \mathbf{v}) = \iint h(\mathbf{u}, \mathbf{v}) \rho(\mathbf{u}) \rho(\mathbf{v}) g(\mathbf{u}, \mathbf{v}) \mathrm{d}\mathbf{u} \mathrm{d}\mathbf{v}$$

Examples of pair correlation functions

In this talk we assume g is isotropic $g(\mathbf{v}, \mathbf{u}) = g(\|\mathbf{v} - \mathbf{u}\|)$.

Examples of g for various point process models:



Reasons for estimating g:

- key summary of clustering/repulsion properties
- if interest focused on parametric model for intensity function

$$\rho(u;\beta) = \exp(z(\mathbf{u})^{\mathsf{T}}\beta)$$

then g needed to evaluate standard errors of parameter estimates $\hat{\beta}$

Kernel estimate

Suppose $\mathbf{X}$ observed in bounded window W.

Kernel estimate ($k_b(\cdot)$ kernel with bandwidth b):

$$\hat{g}(t) = \frac{1}{2\pi t} \sum_{\mathbf{u},\mathbf{v} \in \mathbf{X} \cap W}^{\neq} \frac{k_b(t - \|\mathbf{v} - \mathbf{u}\|)}{\rho(\mathbf{u})\rho(\mathbf{v})|W \cap W_{\mathbf{v}-\mathbf{u}}|}, \quad t > 0$$

or

$$\hat{g}(t) = \frac{1}{2\pi} \sum_{\mathbf{u},\mathbf{v} \in \mathbf{X} \cap W}^{\neq} \frac{k_b(t - \|\mathbf{v} - \mathbf{u}\|)}{\|\mathbf{v} - \mathbf{u}\|\rho(\mathbf{u})\rho(\mathbf{v})|W \cap W_{\mathbf{v}-\mathbf{u}}|}, \quad t > 0$$

Related to non-parametric estimation of probability densities for "distance observations" $D_{uv} = \|\mathbf{v} - \mathbf{u}\|$, $\mathbf{v}, \mathbf{u} \in \mathbf{X}$.

However D_{uv}'s neither independent nor identically distributed. Need to correct for inhomogeneous intensity and edge effects.

Application of Campbell

$$
\begin{aligned}
&\mathbb{E}\hat{g}(t) \\
&= \frac{1}{2\pi}\mathbb{E}\sum_{\mathbf{u},\mathbf{v}\in\mathbf{X}\cap W}^{\neq} \frac{k_b(t-\|\mathbf{v}-\mathbf{u}\|)}{\|\mathbf{v}-\mathbf{u}\|\rho(\mathbf{u})\rho(\mathbf{v})|W\cap W_{\mathbf{v}-\mathbf{u}}|} \\
&= \frac{1}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} 1[\mathbf{u}\in W,\mathbf{v}\in W]\frac{k_b(t-\|\mathbf{v}-\mathbf{u}\|)\rho(\mathbf{u})\rho(\mathbf{v})g(\|\mathbf{v}-\mathbf{u}\|)}{\|\mathbf{v}-\mathbf{u}\|\rho(\mathbf{u})\rho(\mathbf{v})|W\cap W_{\mathbf{v}-\mathbf{u}}|}\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{v} \\
&= \frac{1}{2\pi}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} 1[\mathbf{u}\in W\cap W_{-\mathbf{h}}]\frac{k_b(t-\|\mathbf{h}\|)}{\|\mathbf{h}\||W\cap W_{\mathbf{h}}|}g(\|\mathbf{h}\|)\mathrm{d}\mathbf{u}\mathrm{d}\mathbf{h} \\
&= \int_0^\infty k_b(t-r)g(r)\mathrm{d}r \approx g(t)
\end{aligned}
$$

Choice of band width - minimize MISE

Mean integrated squared error:

$$\mathrm{MISE}(\hat{g}) = \int_0^R \mathbb{E}\big[\hat{g}(r) - g(r)\big]^2 r\mathrm{d}r$$

Minimize estimate of MISE (Guan 2007):

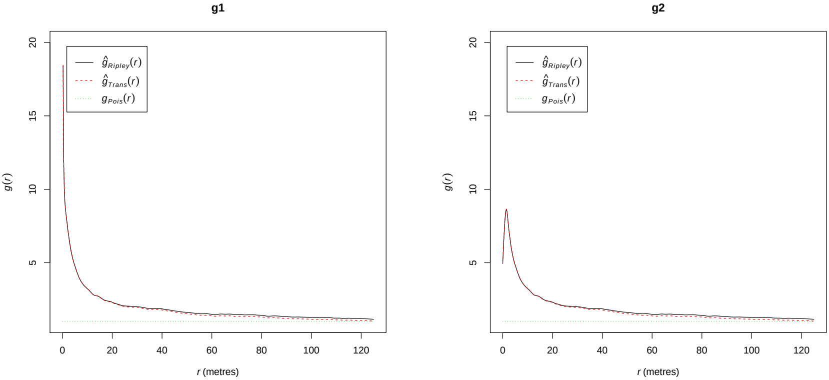
$$\mathrm{MISE}(\hat{g}(\cdot;b)) \approx \int_0^R \hat{g}(r;b)^2 r\mathrm{d}r - 2 \sum_{\substack{\mathbf{u},\mathbf{v}\in X_W \\ \|\mathbf{v}-\mathbf{u}\|\leq R}}^{\neq} \frac{\hat{g}^{-\{\mathbf{u},\mathbf{v}\}}(\|\mathbf{v}-\mathbf{u}\|;b)}{2\pi\rho(\mathbf{u})\rho(\mathbf{v})|W \cap W_{\mathbf{v}-\mathbf{u}}|}$$

Here $\hat{g}^{-\{u,v\}}(\cdot;b)$ is leave one pair out estimate based on $\mathbf{X} \setminus \{u,v\}$ and using band width b.

Bias of kernel estimate

Problem: kernel estimate biased for lags r close to zero.

Two types of kernel estimates for a clustered point pattern:



One can show that left estimate is biased upwards while right estimate is biased downwards.

Bias problem seems difficult to fix:

- Snethlage (2000) suggested adaptive band width to reduce bias but at the expense of large variance
- Guan (2007) derived correction near zero but assuming Poisson process

Orthogonal series estimation

Orthogonal series estimate (OSE) of probability density on interval I:

$$f(x) = \sum_{k=1}^{\infty} \theta_k \phi_k(x) \quad \theta_k = \int_I f(z)\phi_k(z)w(z)\mathrm{d}z$$

where $\{\phi_k\}_k$ orthogonal basis for weight function $w(\cdot)$.

Suppose $X_1, \ldots, X_n$ *iid* sample from f. Then

$$\hat{\theta}_k = \frac{1}{n}\sum_{i=1}^{n} \phi_k(X_i)w(X_i)$$

unbiased estimate of θ_k.

In practice truncated (hence biased) estimate is used

$$\hat{f}(x) = \sum_{k=1}^{K} \hat{\theta}_k \phi_k(x)$$

OSE for pair correlation function

Estimate g on interval $]0, R[$:

$$g(r) = \sum_{k=1}^{\infty} \theta_k \phi_k(r) \quad \theta_k = \int_0^R g(r)\phi_k(r)w(r)\mathrm{d}r$$

By Campbell formula we obtain unbiased estimate

$$\hat{\theta}_k = \frac{1}{2\pi} \sum_{\substack{\mathbf{u},\mathbf{v}\in\mathbf{X}_W: \\ \|\mathbf{u}-\mathbf{v}\|\leq R}}^{\neq} \frac{\phi_k(\|\mathbf{v}-\mathbf{u}\|)w(\|\mathbf{v}-\mathbf{u}\|)}{\|\mathbf{v}-\mathbf{u}\|\rho(\mathbf{u})\rho(\mathbf{v})|W \cap W_{\mathbf{v}-\mathbf{u}}|}$$

Similar to estimate for probability density estimation but

- $D_{uv} = \|\mathbf{v}-\mathbf{u}\|$ neither independent nor identically distributed
- we have to correct for spatially varying intensity and edge effects using factor $\rho(\mathbf{u})\rho(\mathbf{v})|W \cap W_{\mathbf{v}-\mathbf{u}}|$

Choices of basis

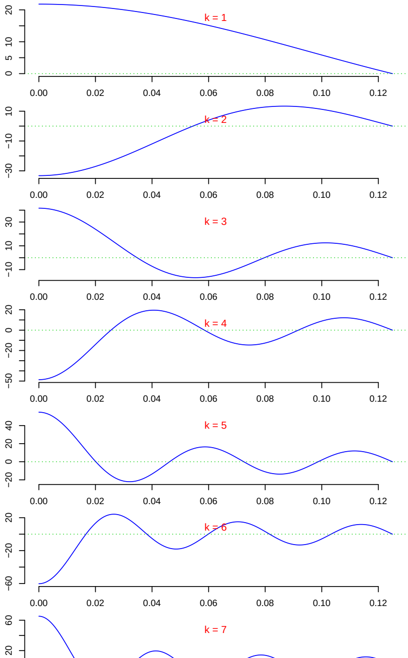
Cosine ($w(r) = 1$):

$$\phi_1(r) = \frac{1}{\sqrt{R}}, \quad \phi_k(r) = \frac{\sqrt{2}}{\sqrt{R}} \cos\left((k-1)\pi r/R\right), \ k \geq 2$$

Bessel ($w(r) = r$):

$$\phi_k(r) = \frac{\sqrt{2}}{R J_1(\alpha_{0,k})} J_0\left(r \frac{\alpha_{0,k}}{R}\right), \ k \geq 1$$

Bessel basis



In practice we use truncated estimate

$$\hat{g}(r) = \sum_{k=1}^{K} \hat{\theta}_k \phi_k(r)$$

How to choose K ?

(problem analogous to band width selection for kernel estimate)

Option: choose K that minimizes estimate of mean integrated squared error

Mean integrated squared error (MISE)

$$\mathrm{MISE}(\hat{g}) = \int_0^R \mathbb{E}\big[\hat{g}(r) - g(r)\big]^2 w(r)\mathrm{d}r$$

For OSE,

$$\mathrm{MISE}\big(\hat{g}\big) \equiv \sum_{k=1}^{K} \Big[\mathbb{E}(\hat{\theta}_k)^2 - 2\theta_k^2\Big]$$

Idea: replace $\mathbb{E}(\hat{\theta}_k)^2$ by $(\hat{\theta}_k)^2$ and replace θ_k^2 by (asymptotically) unbiased estimate $\widehat{\theta_k^2} =$

$$\frac{1}{4\pi^2} \sum_{\substack{\mathbf{u},\mathbf{v},\mathbf{u}',\mathbf{v}' \in \mathbf{X} \cap W: \\ \|\mathbf{v}-\mathbf{u}\| \le R \\ \|\mathbf{v}'-\mathbf{u}'\| \le R}}^{\neq} \frac{\phi_k(\|\mathbf{v}-\mathbf{u}\|)\phi_k(\|\mathbf{v}'-\mathbf{u}'\|)w(\|\mathbf{v}-\mathbf{u}\|)w(\|\mathbf{v}'-\mathbf{u}'\|)}{\|\mathbf{v}-\mathbf{u}\|\|\mathbf{v}'-\mathbf{u}'\|\rho(\mathbf{u})\rho(\mathbf{v})\rho(\mathbf{u}')\rho(\mathbf{v}')|W \cap W_{\mathbf{v}-\mathbf{u}}||W \cap W_{\mathbf{v}'-\mathbf{u}'}|}$$

Choose K to minimize estimate of $\mathrm{MISE}(\hat{g})$.

Simple smoothing scheme:

$$\hat{g}(r) = \sum_{k=1}^{K} \hat{\theta}_k \phi_k(r)$$

Refined smoothing scheme:

$$\hat{g}(r) = \sum_{k=1}^{K} \frac{\widehat{\theta_k^2}}{(\hat{\theta}_k)^2} \hat{\theta}_k \phi_k(r)$$

Wahba smooothing scheme:

$$\hat{g}(r) = \sum_{k=1}^{K} \frac{1}{1 + c_1 k^{c_2}} \hat{\theta}_k \phi_k(r)$$

Asymptotic results

Assume $\mathbf{X}$ observed on increasing sequence of windows W_n.

$\hat{g}_n$ is estimate based on $\mathbf{X} \cap W_n$ with truncation K_n.

Need a number of 'standard' conditions on point process and smoothing scheme (e.g. increasing K_n but $K_n/|W_n| \to 0$).

Asymptotic results: consistency

Crucial assumption (since we divide by $\|\mathbf{v} - \mathbf{u}\|$ in $\hat{\theta}_k$):

$$g(r)\frac{w(r)}{r} \leq C$$

to ensure finite variance $\mathbb{V}\mathrm{ar}\hat{\theta}_{k,n}$.

This is OK for Bessel basis ($w(r) = r$) if g bounded.

For cosine basis ($w(r) = 1$) we need $g(r) \to 0$ as $r \to 0$ (e.g. determinantal point process).

We then have consistency in MISE:

$$\mathrm{MISE}(\hat{g}_n) = \int_0^R \mathbb{E}\big[\hat{g}_n(r) - g(r)\big]^2 w(r)\mathrm{d}r \to 0$$

Asymptotic results: normality

Under suitable (α) mixing and moment conditions we have CLT (Biscio and Waagepetersen, 2017) for spatial averages of the form

$$S_n = \frac{1}{|W_n|} \sum_{\substack{\mathbf{u},\mathbf{v} \in \mathbf{X} \cap W_n \\ \|\mathbf{v}-\mathbf{u}\| \leq R}}^{\neq} \frac{f_n(\mathbf{v}-\mathbf{u})}{\rho(\mathbf{u})\rho(\mathbf{v}) e_n(\mathbf{v}-\mathbf{u})} \quad e_n(\mathbf{v}-\mathbf{u}) = \frac{|W_n|}{|W_n \cap W_{n,\mathbf{v}-\mathbf{u}}|}$$

whenever f_n is bounded.

This immediately gives CLT for $\hat{\theta}_{k,n}$.

The OSE of g can be rewritten as

$$\hat{g}_n(r) = \frac{1}{2\pi|W_n|} \sum_{\substack{\mathbf{u},\mathbf{v}\in\mathbf{X}\cap W_n: \\ \mathbf{u}-\mathbf{v}\in B_R}}^{\neq} \frac{w(\|\mathbf{v}-\mathbf{u}\|)h_n(\mathbf{v}-\mathbf{u}, r)}{\rho(\mathbf{u})\rho(\mathbf{v})\|\mathbf{v}-\mathbf{u}\|e_n(\mathbf{v}-\mathbf{u})|}$$

where

$$h_n(\mathbf{h}, r) = \sum_{k=1}^{K_n} \phi_k(\|\mathbf{h}\|)\phi_k(r).$$

Here

$$f_n(\mathbf{h}) = \frac{w(\mathbf{h})h_n(\mathbf{h})}{\|\mathbf{h}\|}$$

is not bounded but we assume h_n is $O(K_n^\omega)$ for some $\omega > 0$. Then we have

$$[\mathbb{V}\mathrm{ar}\hat{g}_n(r)]^{-1/2}\left(\hat{g}_n(r) - \mathbb{E}\hat{g}_n(r)\right) \xrightarrow{D} N(0,1).$$

provided

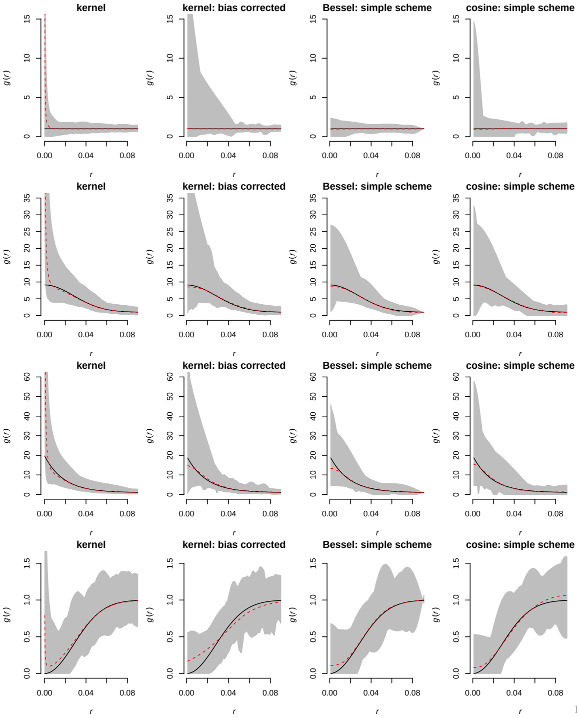
$$\liminf_{n\to\infty} \frac{\mathbb{V}\mathrm{ar}\hat{g}_n(r)}{K_n^{2\omega}|W_n|} > 0$$

Conclusions from simulation studies

We conducted simulation studies for Poisson, Thomas and VarGamma processes (clustered) and determinantal processes (repulsive).

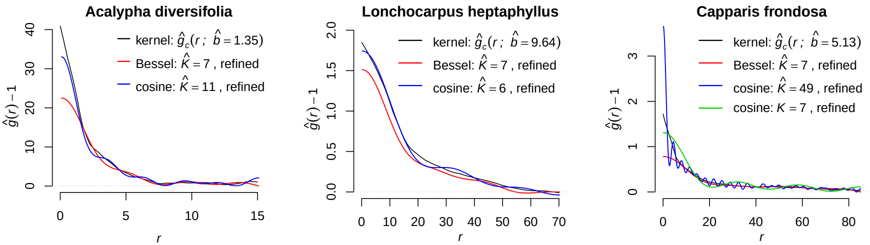
Poisson and cluster processes: with Bessel basis substantial improvement in MISE relative to kernel estimates (cosine as good as or better than kernel)

Determinantal: OSE with cosine basis or Bessel with simple smoothing scheme works as well as kernel estimates.



kernel
kernel: bias corrected
Bessel: simple scheme
cosine: simple scheme
kernel
kernel: bias corrected
Bessel: simple scheme
cosine: simple scheme
kernel
kernel: bias corrected
Bessel: simple scheme
cosine: simple scheme
kernel
kernel: bias corrected
Bessel: simple scheme
cosine: simple scheme
g(r)
r
0.00
0.04
0.08
0
5
10
15
20
25
30
35
40
50
60
0.0
0.5
1.0
1.5

Application to tropical tree locations



Future work

Disadvantage: global basis functions - estimates depends on chosen upper range R.

- Wavelet basis ?
- Frames ?

References

Guan, Y. (2007) A least-squares cross-validation approach in pair correlation function estimations, *Statistics and Probability Letters*, **77**, 1722-1729.

Jalilian, A., Guan, Y. and Waagepetersen, R. (2017) Orthogonal series estimation of the pair correlation function of a spatial point process, under revision.

Waagepetersen, R. and Jalilian, A. (2017) Fast band width selection for estimation of the pair correlation function, submitted.

Thanks for your attention !