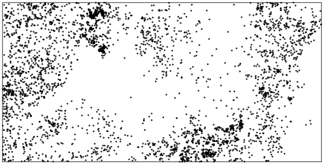
Two-step estimation for inhomogeneous spatial point processes

Rasmus Waagepetersen
Department of Mathematics
Aalborg University
Denmark

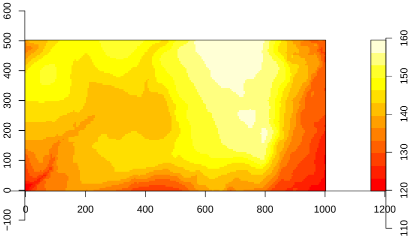
November 19, 2007

Tropical rain forests trees

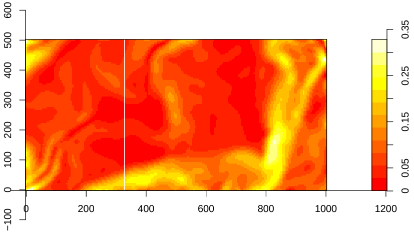
Beilschmiedia



- *observation window* $= 1000 \text{ m} \times 500 \text{ m}$
- seed dispersal $\Rightarrow$ *clustering*
- *covariates* $\Rightarrow$ *inhomogeneity*



Elevation



Norm of elevation gradient (slope)

Intensity function and product density

Intensity function of point process **X** on $\mathbb{R}^2$:

$$\rho(u)\mathrm{d}A \approx P(\textbf{X} \text{ has a point in A})$$

Intensity function and product density

Intensity function of point process $\mathbf{X}$ on $\mathbb{R}^2$:

$$\rho(u)\mathrm{d}A \approx P(\mathbf{X} \text{ has a point in A})$$

Second order product density

$$\rho^{(2)}(u, v)dAdB \approx P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)$$

Intensity function and product density

Intensity function of point process $\mathbf{X}$ on $\mathbb{R}^2$:

$$\rho(u)\mathrm{d}A \approx P(\mathbf{X} \text{ has a point in A})$$

Second order product density

$$\rho^{(2)}(u, v)dAdB \approx P(\mathbf{X} \text{ has a point in each of } A \text{ and } B)$$

Pair correlation and K-function (provided $g(u, v) = g(u - v)$)

$$g(u, v) = \frac{\rho^{(2)}(u, v)}{\rho(u)\rho(v)} \quad \text{and } K(t) = \int_{\mathbb{R}^2} 1[\|u\| \leq t]g(u)\mathrm{d}u$$

NB: for Poisson process, $g(u - v) = 1$, clustering: $g(u - v) > 1$.

Parametric models

Study influence of covariates using log-linear model for intensity function:

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

and quantify clustering using parameter ψ in parametric model

$$K(t;\psi) = \int_{\|v\|\leq t} g(v;\psi)\mathrm{d}v$$

for K/g-function.

Estimating function for β

Maximum likelihood estimation only easy in case of a Poisson process $\mathbf{X}$ in which case log likelihood is

$$l(\beta) = \sum_{u \in \mathbf{X} \cap W} z(u)\beta^{\mathsf{T}} - \int_W \rho(u;\beta)\mathrm{d}u$$

Poisson score estimating function based on point process $\mathbf{X}$ observed in W:

$$u_1(\beta) = \sum_{u \in \mathbf{X} \cap W} z(u) - \int_W z(u)\rho(u;\beta)\mathrm{d}u$$

also applicable for *non-Poisson* point processes with intensity function $\rho(\cdot;\beta)$ (Schoenberg, 2004, Waagepetersen, 2007)

Estimating function for ψ

Estimate of K-function:

$$\hat{K}_\beta(t) = \sum_{u,v \in \mathbf{X} \cap W} \frac{1[0 < \|u - v\| \leq t]}{\rho(u;\beta)\rho(v;\beta)|W \cap W_{u-v}|}$$

Unbiased if $\beta = \beta^*$ 'true' regression parameter.

Minimum contrast estimation: minimize

$$\int_0^r \left(\hat{K}_\beta(t) - K(t;\psi)\right)^2 \mathrm{d}t$$

or solve estimating equation

$$u_{2,\beta}(\psi) = |W| \int_0^r \left(\hat{K}_\beta(t) - K(t;\psi)\right) \frac{\mathrm{d}K(t;\psi)}{\mathrm{d}\psi} \mathrm{d}t = 0$$

Two-step estimation

Estimate $(\hat{\beta}, \hat{\psi})$ by solving

1. $u_1(\beta) = 0$
2. $u_{2,\hat{\beta}}(\psi) = 0$

or equivalently solve

$$u(\beta, \psi) = \big(u_1(\beta), u_{2,\beta}(\psi)\big) = 0.$$

Waagepetersen and Guan (2007): asymptotic properties of $(\hat{\beta}, \hat{\psi})$.

CLT for estimating function

Consider increasing observation windows W_n.

Divide $\mathbb{R}^2$ into quadratic cells $A_{ij} = s[i, i+1[\times s[j, j+1[$ of area s^2.

Express Poisson score in terms of lattice process X_{ij}, $i, j \in \mathbb{Z}$:

$$u_1^n(\beta) = \sum_{u \in \mathbf{X} \cap W_n} z(u) - \int_{W_n} z(u)\rho(u;\beta)\mathrm{d}u =$$

$$\sum_{i,j} \left[\sum_{u \in \mathbf{X} \cap W_n \cap A_{ij}} z(u) - \int_{W_n \cap A_{ij}} z(u)\rho(u;\beta)\mathrm{d}u \right] = \sum_{ij : A_{ij} \subseteq W_n} X_{ij} + o_P(1)$$

Similarly:

$$u^n_{2,\beta}(\psi) = |W_n| \int_0^r \left(\hat{K}_\beta(t) - K(t;\psi)\right) \frac{\mathrm{d}K(t;\psi)}{\mathrm{d}\psi} \mathrm{d}t = \sum_{ij: A_{ij} \subseteq W_n} Y_{ij} + o_P(1)$$

where

$$Y_{ij} = \int_0^r s^2(\hat{K}_{\beta,ij}(t) - K(t;\psi)) \frac{\mathrm{d}K(t;\psi)}{\mathrm{d}\psi} \mathrm{d}t$$

and

$$\hat{K}_{\beta,ij}(t) = \frac{1}{s^2} \sum_{u \in \mathbf{X} \cap A_{ij}, v \in \mathbf{X}} \frac{1[0 < \|u - v\| \leq t]}{\rho(u;\beta)\rho(v;\beta)}$$

estimate of K-function based on $\mathbf{X} \cap A_{ij} \oplus r$.

Apply Guyon/Bolthausen CLT for mixing lattice processes to random field $\{Z_{ij}\}_{ij}$ of linear combinations

$$Z_{ij} = X_{ij} x^{\mathsf{T}} + Y_{ij} y^{\mathsf{T}}.$$

Use result in Aitchison and Silvey (1958) to show that there exist $O_P(|W_n|^{-1/2})$ consistent sequences of solutions $\hat{\beta}_n$ and $\hat{\psi}_n$ of $u_1^n(\beta) = 0$ and $u^n_{2,\hat{\beta}_n}(\psi) = 0$.

Finally, Taylor expansion:

$$|W_n|^{-1/2} u^n(\beta^*, \psi^*) = |W_n|^{1/2}[(\hat{\beta}_n, \hat{\psi}_n) - (\beta^*, \psi^*)] \frac{J_n(\tilde{\beta}, \tilde{\psi})}{|W_n|}$$

where

$$J_n(\beta, \psi) = \begin{bmatrix} \frac{\mathrm{d}}{\mathrm{d}\beta^\mathsf{T}} u_{n,1}(\beta) & \frac{\mathrm{d}}{\mathrm{d}\beta^\mathsf{T}} u_{n,2}(\beta, \psi) \\ 0 & \frac{\mathrm{d}}{\mathrm{d}\psi^\mathsf{T}} u_{n,2}(\beta, \psi) \end{bmatrix}$$

and $\frac{J_n(\tilde{\beta}, \tilde{\psi})}{|W_n|} - I_n \to 0$ for non-random matrices I_n. Hence

$$|W_n|^{1/2}[(\hat{\beta}_n, \hat{\psi}_n) - (\beta^*, \psi^*)] I_n \Sigma_n^{-1/2} \to N(0, I)$$

where Σ_n variance matrix for $|W_n|^{-1/2} u^n(\beta^*, \psi^*)$.

Mixing

Consider $E_1, E_2 \subseteq \mathbb{R}^2$ point configurations F_1 and F_2.

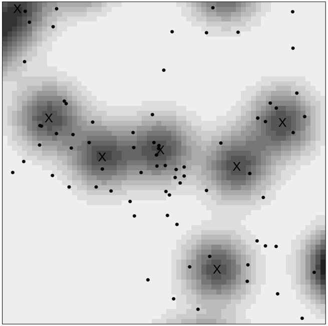
Need polynomial decay of

$$|P(\mathbf{X} \cap E_1 \in F_1, \mathbf{X} \cap E_2 \in F_2) - P(\mathbf{X} \cap E_1 \in F_1)P(\mathbf{X} \cap E_2 \in F_2)|$$

as function of distance between E_1 and E_2.

This can easily be verified for a Poisson cluster process where cluster density decays fast enough.

Example: modified Thomas process



Mothers (crosses) stationary Poisson point process $\mathbf{M}$ with intensity $\kappa > 0$.

Clusters $\mathbf{X}_m, m \in \mathbf{M}$ Poisson processes of offspring dispersed according to $k =$ bivariate isotropic Gaussian density.

ω: standard deviation of Gaussian density

α: Expected number of offspring for each mother.

Cox process with random intensity function:

$$\Lambda(u) = \alpha \sum_{m \in \mathbf{M}} k(u - m; \omega)$$

Inhomogeneous Thomas process

$z_{1:p}(u) = (z_1(u), \dots, z_p(u))$ vector of p nonconstant covariates.

$\beta_{1:p} = (\beta_1, \dots \beta_p)$ regression parameter.

Inhomogeneous random intensity function:

$$\Lambda_{\text{inhom}}(u) = \exp(z(u)_{1:p}\beta_{1:p}^{\mathsf{T}})\Lambda(u)$$

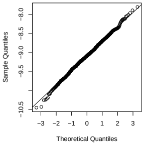
$$\rho(u;\beta) = \exp(z(u)\beta^{\mathsf{T}}) \quad \beta = (\beta_0, \beta_1, \dots, \beta_p) = (\log\kappa\alpha, \beta_1, \dots, \beta_p)$$
$$\psi = (\kappa, \omega)$$

Thomas process mixing and inhomogenous Thomas process independent thinning of Thomas $\Rightarrow$ inhomogeneous Thomas mixing too.

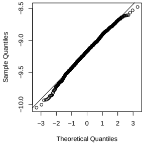
Simulations: $\log \hat{\kappa}$

QQ-plots with varying expected numbers of mothers/offspring.

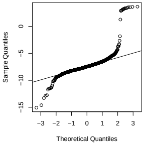
50/200



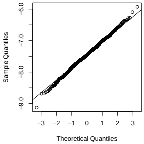
50/800



250/200



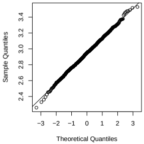
250/800



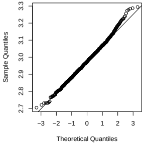
Simulations: $\log \hat{\omega}$

QQ-plots with varying expected numbers of mothers/offspring.

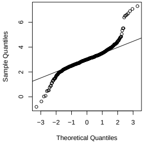
50/200



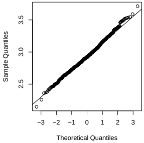
50/800



250/200



250/800



Tropical trees

Estimates of κ and ω for model with altitude and gradient: 8×10^{-5} and 20.

Estimates of κ and ω for model with altitude and gradient *and* soil variables nitrogen, phosphorous, potassium, pH: 2.2×10^{-4} and 13.

Hence much residual clustering explained by soil variables.

Confidence intervals: $[1.2 \times 10^{-4}, 3.9 \times 10^{-4}]$ and [10,17].

Issues

- choice of integration limit r for minimum contrast estimation

$$\int_0^r \left(\hat{K}_{\hat{\beta}}(t) - K(t;\psi)\right)^2 \mathrm{d}t$$

- variance of $\hat{K}_{\hat{\beta}}(t)$ smaller than variance of $\hat{K}_{\beta^*}(t)$ hence better to use $\hat{\beta}$ than β^* when estimating ψ.
- LGCPs mixing if Gaussian field is mixing - but only mixing results for Gaussian lattice processes.
- two-step estimation only depends on first and second order properties - but when is a parametric model $K(\cdot;\psi)$ a legitimate K-function ?