Minimum contrast estimation for inhomogeneous space-time cluster point processes

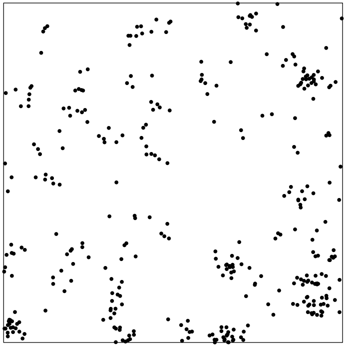
Jiří Dvořák[1,2], joint work with Michaela Prokešová[1]

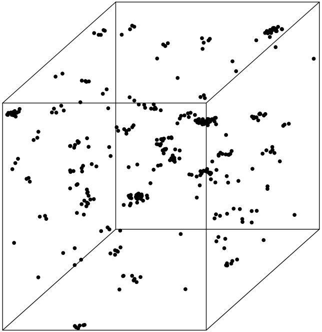
[1]Faculty of Mathematics and Physics, Charles University in Prague,
[2]Institute of Information Theory and Automation of the CAS, Prague

Spatial Statistics and Image Analysis in Biology

26. 5. 2016

Point processes – spatial and space-time





Stationary spatial cluster point process

Cox process model with

$$\Lambda(u) = \sum_{v \in \Phi} \mu\, k(v - u), \quad u \in \mathbb{R}^2$$

Φ – Poisson p. p. in $\mathbb{R}^2$ with intensity $\kappa > 0$

μ – mean number of offsprings per parent point

k – smoothing kernel (p.d.f. on $\mathbb{R}^2$)

Stationary spatial cluster point process

Cox process model with

$$\Lambda(u) = \sum_{v \in \Phi} \mu\, k(v - u), \quad u \in \mathbb{R}^2$$

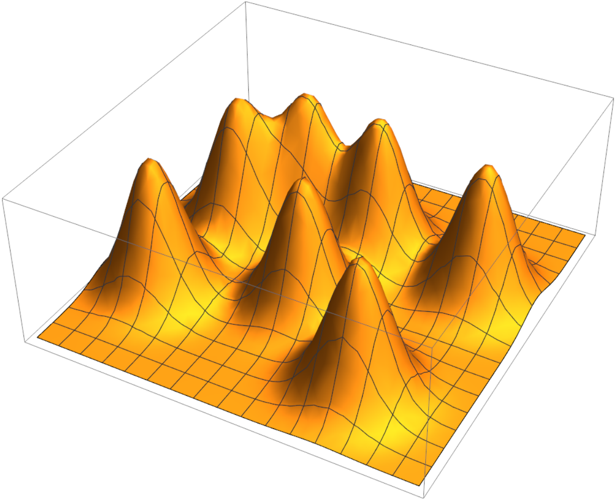
Φ – Poisson p. p. in $\mathbb{R}^2$ with intensity $\kappa > 0$

μ – mean number of offsprings per parent point

k – smoothing kernel (p.d.f. on $\mathbb{R}^2$)

(Poisson-Neyman-Scott model; similarly for shot-noise Cox processes)

Stationary spatial cluster point process



Stationary space-time cluster point process

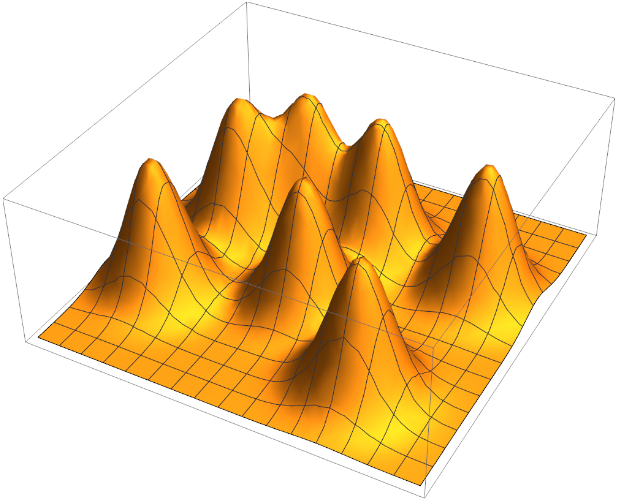
$$\Lambda(u,t) = \sum_{(v,s)\in\Phi} \mu\, k(v-u, s-t), \quad (u,t) \in \mathbb{R}^2 \times \mathbb{R}$$

Φ – Poisson p. p. on $\mathbb{R}^2 \times \mathbb{R}$ with intensity κ

μ – mean number of offsprings per parent point

k – smoothing kernel (p.d.f. on $\mathbb{R}^2 \times \mathbb{R}$)

Stationary space-time cluster point process



SOIRS point process

Instead of stationarity we assume *second-order intensity reweighted stationarity:*

SOIRS point process

Instead of stationarity we assume *second-order intensity reweighted stationarity:*

$$g((u,t),(v,s)) = \frac{\lambda_2((u,t),(v,s))}{\lambda(u,t)\lambda(v,s)} = g(v-u, s-t)$$

$$K(r,t) = \int_{\|s\|<r} \int_{|\tau|<t} g(s,\tau)\,\mathrm{d}\tau\,\mathrm{d}s$$

SOIRS point process

Instead of stationarity we assume *second-order intensity reweighted stationarity:*

$$g((u,t),(v,s)) = \frac{\lambda_2((u,t),(v,s))}{\lambda(u,t)\lambda(v,s)} = g(v-u, s-t)$$

$$K(r,t) = \int_{\|s\|<r} \int_{|\tau|<t} g(s,\tau)\,\mathrm{d}\tau\,\mathrm{d}s$$

Example: location-dependent thinning of a stationary process

Inhomogeneous space-time cluster point process

- inhomogeneity introduced by location-dependent thinning
- $0 \leq f \leq 1 \ldots$ „inhomogeneity function“
- each point (u, t) of a stationary process retained with probability $f(u, t)$, independently of other points
- f may depend on a vector of covariates $z(u, t)$, e.g. $f(u, t) \propto \exp\left\{\beta^T z(u, t)\right\}$

Notation and model parametrization

- X ... thinned (inhomogeneous) space-time process
- λ ... its (first order) intensity function
- $W \times T$ ... observation window, $|W| > 0, |T| > 0$
- k ... smoothing kernel (p.d.f. on $\mathbb{R}^2 \times \mathbb{R}$)
- $\max_{W \times T} f = 1$

Notation and model parametrization

- X ... thinned (inhomogeneous) space-time process
- λ ... its (first order) intensity function
- $W \times T$ ... observation window, $|W| > 0, |T| > 0$
- k ... smoothing kernel (p.d.f. on $\mathbb{R}^2 \times \mathbb{R}$)
- $\max_{W \times T} f = 1$

- β ... (vector) parameter of the inhom. function $f(\cdot; \beta)$
- ψ ... (vector) parameter of the smoothing kernel $k(\cdot, \cdot; \psi)$
- κ, μ ... intensity of parents, mean number of offsprings per parent

Step-wise parameter estimation

MLE computationally demanding, an alternative approach (inspired by Waagepetersen, Guan, 2009):

Step-wise parameter estimation

MLE computationally demanding, an alternative approach (inspired by Waagepetersen, Guan, 2009):

1. estimate the inhomogeneity parameters β by maximizing the Poisson likelihood function

Step-wise parameter estimation

MLE computationally demanding, an alternative approach (inspired by Waagepetersen, Guan, 2009):

1. estimate the inhomogeneity parameters β by maximizing the Poisson likelihood function
2. estimate the interaction parameters ψ using minimum contrast on the space-time K-function

Step-wise parameter estimation

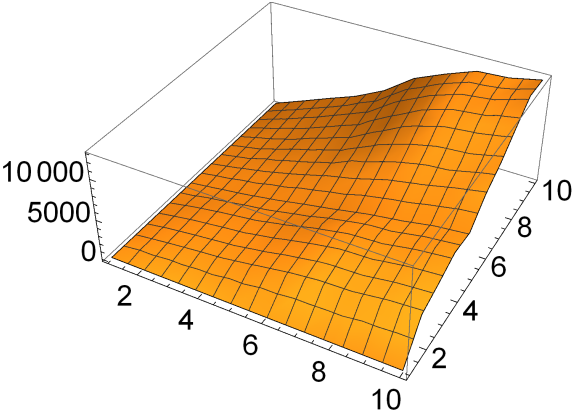
MLE computationally demanding, an alternative approach (inspired by Waagepetersen, Guan, 2009):

1. estimate the inhomogeneity parameters β by maximizing the Poisson likelihood function
2. estimate the interaction parameters ψ using minimum contrast on the space-time K-function
3. estimate the remaining parameters (κ, μ) from the previous estimates and the total intensity of points

Second step – minimum contrast estimation

Beneš et al. (2015), “full minimum contrast”

$$\widehat{\psi} = \arg\min \int_{r_{min}}^{r_{max}} \int_{t_{min}}^{t_{max}} (\widehat{K}(r,t)^q - K(r,t;\psi)^q)^2 \,\mathrm{d}t\,\mathrm{d}r$$



Second step – minimum contrast estimation

$$\widehat{K}(r,t) =$$

$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j) \in X \cap (W \times T)}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j) w_2(t_i,t_j) \lambda(u_i,t_i;\widehat{\beta}) \lambda(u_j,t_j;\widehat{\beta})}$$

Second step – minimum contrast estimation

$$\widehat{K}(r,t) =$$
$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j)\in X\cap(W\times T)}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j)w_2(t_i,t_j)\lambda(u_i,t_i;\widehat{\beta})\lambda(u_j,t_j;\widehat{\beta})}$$

Properties:

+ does not need space-time separability of $\lambda(u,t), f(u,t)$

Second step – minimum contrast estimation

$$\widehat{K}(r,t) =$$

$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j)\in X\cap(W\times T)}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j)w_2(t_i,t_j)\lambda(u_i,t_i;\widehat{\beta})\lambda(u_j,t_j;\widehat{\beta})}$$

Properties:

+ does not need space-time separability of $\lambda(u,t), f(u,t)$
- optimizing w.r.t. all the interaction parameters at once

Second step – minimum contrast estimation

$$\widehat{K}(r,t) =$$
$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j) \in X \cap (W \times T)}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j) w_2(t_i,t_j) \lambda(u_i,t_i;\widehat{\beta}) \lambda(u_j,t_j;\widehat{\beta})}$$

Properties:

+ does not need space-time separability of $\lambda(u,t), f(u,t)$
- optimizing w.r.t. all the interaction parameters at once
- requires fine grid $(r_i, t_i) \Rightarrow$ computational burden

Second step – minimum contrast estimation

$$\widehat{K}(r,t) =$$
$$\frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j) \in X \cap (W \times T)}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j) w_2(t_i,t_j) \lambda(u_i,t_i;\widehat{\beta}) \lambda(u_j,t_j;\widehat{\beta})}$$

Properties:

+ does not need space-time separability of $\lambda(u,t), f(u,t)$
- optimizing w.r.t. all the interaction parameters at once
- requires fine grid $(r_i, t_i) \Rightarrow$ computational burden
- sensitive to the choice of initial estimates

$$\widehat{K}(r,t) = \frac{1}{|W \times T|} \sum_{(u_i,t_i),(u_j,t_j) \in X \cap (W \times T)}^{\neq} \frac{I(\|u_i - u_j\| \leq r, |t_i - t_j| \leq t)}{w_1(u_i,u_j) w_2(t_i,t_j) \lambda(u_i,t_i;\widehat{\beta}) \lambda(u_j,t_j;\widehat{\beta})}$$

Properties:

- **+** does not need space-time separability of $\lambda(u,t), f(u,t)$
- **-** optimizing w.r.t. all the interaction parameters at once
- **-** requires fine grid $(r_i, t_i) \Rightarrow$ computational burden
- **-** sensitive to the choice of initial estimates

Can we have the **+** without the **-**?

Profile contrast estimation

Assume (Møller, Ghorbani, 2012):

$k(w, \tau; \psi) = k_1(w; \psi_1) k_2(\tau; \psi_2), \; w \in \mathbb{R}^2, \tau \in \mathbb{R}$

Profile contrast estimation

Assume (Møller, Ghorbani, 2012):

$$k(w, \tau; \psi) = k_1(w; \psi_1) k_2(\tau; \psi_2), \; w \in \mathbb{R}^2, \tau \in \mathbb{R}$$

$$K(r, t) = 2\pi r^2 t + \frac{1}{\kappa} \cdot P_1(r; \psi_1) \cdot P_2(t; \psi_2), \; r \geq 0, t \geq 0$$

Profile contrast estimation

Assume (Møller, Ghorbani, 2012):

$$k(w, \tau; \psi) = k_1(w; \psi_1)k_2(\tau; \psi_2),\ w \in \mathbb{R}^2, \tau \in \mathbb{R}$$

$$K(r, t) = 2\pi r^2 t + \frac{1}{\kappa} \cdot P_1(r; \psi_1) \cdot P_2(t; \psi_2),\ r \geq 0, t \geq 0$$

Even “very non-separable” choices of k (e.g. ambit processes) result in little deviation from this structure

Profile contrast estimation

Assume (Møller, Ghorbani, 2012):

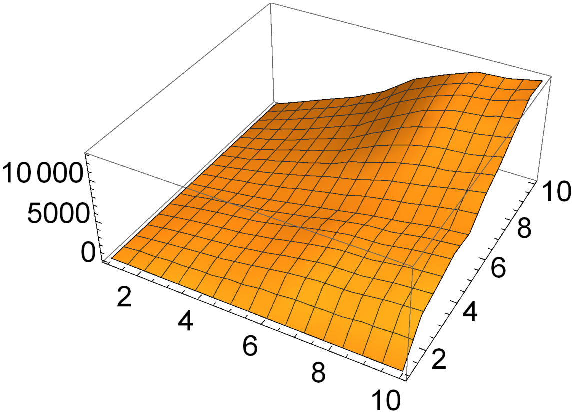
$$k(w, \tau; \psi) = k_1(w; \psi_1)k_2(\tau; \psi_2), \; w \in \mathbb{R}^2, \tau \in \mathbb{R}$$

$$K(r, t) = 2\pi r^2 t + \frac{1}{\kappa} \cdot P_1(r; \psi_1) \cdot P_2(t; \psi_2), \; r \geq 0, t \geq 0$$

Even “very non-separable” choices of k (e.g. ambit processes) result in little deviation from this structure – not possible to see from empirical estimate of $K(r, t)$.

Profile contrast estimation

$$K(r,t) = 2\pi r^2 t + \frac{1}{\kappa} \cdot P_1(r; \psi_1) \cdot P_2(t; \psi_2), \;\; r \geq 0, t \geq 0$$



Profile contrast estimation I.

- choose spatial range $R > 0$
- estimate $\widehat{K}(R, t)$ as a function of t
- use minimum contrast on

$$\begin{aligned} K(R,t) &= 2\pi R^2 t + \frac{1}{\kappa} \cdot P_1(R;\psi_1) \cdot P_2(t;\psi_2) \\ &= 2\pi R^2 t + c(\kappa, \psi_1, R) \cdot P_2(t;\psi_2) \end{aligned}$$

Profile contrast estimation II.

- choose temporal lag $\tau > 0$
- estimate $\widehat{K}(r, \tau)$ as a function of r
- use minimum contrast on

$$\begin{aligned} K(r, \tau) &= 2\pi r^2 \tau + \frac{1}{\kappa} \cdot P_2(\tau; \psi_2) \cdot P_1(r; \psi_1) \\ &= 2\pi r^2 \tau + c(\kappa, \psi_2, \tau) \cdot P_1(r; \psi_1) \end{aligned}$$

Profile contrast estimation – summary

- optimization in lower-dimensional space

Profile contrast estimation – summary

- optimization in lower-dimensional space
- less sensitive to “wrong” choice of initial estimates

Profile contrast estimation – summary

- optimization in lower-dimensional space
- less sensitive to “wrong” choice of initial estimates
- only works with the profiles of the grid $(r_i, \tau), (R, t_i)$

Profile contrast estimation – summary

- optimization in lower-dimensional space
- less sensitive to "wrong" choice of initial estimates
- only works with the profiles of the grid $(r_i, \tau), (R, t_i)$
- consistency and as. normality of $\widehat{\beta}, \widehat{\psi}_1, \widehat{\psi}_2$ under suitable moment and mixing conditions

Profile contrast estimation – summary

- optimization in lower-dimensional space
- less sensitive to “wrong” choice of initial estimates
- only works with the profiles of the grid $(r_i, \tau), (R, t_i)$
- consistency and as. normality of $\widehat{\beta}, \widehat{\psi}_1, \widehat{\psi}_2$ under suitable moment and mixing conditions
- in simulation studies
    - with reasonable initialization, almost as efficient as full minimum contrast

Profile contrast estimation – summary

- optimization in lower-dimensional space
- less sensitive to “wrong” choice of initial estimates
- only works with the profiles of the grid $(r_i, \tau), (R, t_i)$
- consistency and as. normality of $\widehat{\beta}, \widehat{\psi}_1, \widehat{\psi}_2$ under suitable moment and mixing conditions
- in simulation studies
  - with reasonable initialization, almost as efficient as full minimum contrast
  - with bad initialization, much more stable than full minimum contrast

References

1. Beneš, V. et al. (2015). *Space-time models in stochastic geometry.* In: Stochastic Geometry, Spatial Statistics and Random Fields: Models and Algorithms (ed. V. Schmidt). Springer, Heidelberg.
2. Møller, J., Ghorbani, M. (2012). *Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes.* Stat. Neerlandica **66**, 472 – 491.
3. Waagepetersen, R. P., Guan, Y. (2009). *Two-step estimation for inhomogeneous spatial point processes.* J. R. Stat. Soc. B **71**, 685 – 702.