Point patterns on the sphere

Ege Rubak

Collaborators

Main reference is a soon to be published (hopefully!?) paper for special issue of Spatial Statistics together with Jesper Møller.

Lots of technical things related to defining determinantal point processes on the sphere were done in another paper together with Jesper, Morten Nielsen, and Emilio Porcu. It is currently in review (we think) but available as preprint.

Lately I have started to work with Adrian Baddeley, Tom Lawrence, and Tuomas Rajala on merging different packages/tools for point patterns on the sphere (and general geometry on the sphere).

Plan

- Briefly mention some models on the sphere.
- Palm distributions and summary statistics on the sphere.
- Software

Models

- Poisson process
- Thomas type cluster processes (Baddeley and coauthors)
- Determinantal processes (us)

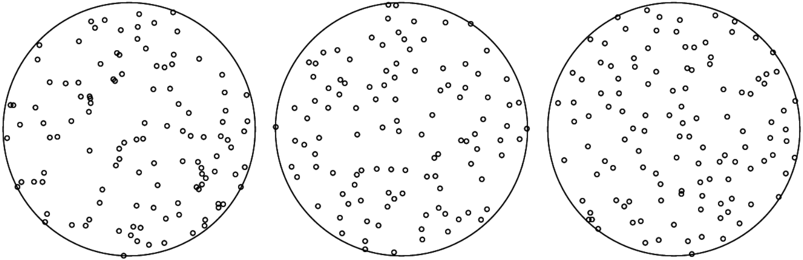
Determinantal point processes

- Specified in terms of kernel (pos. semi def. function) on sphere.
- Kernel determines intensity and higher order moments.
- Requires spectral decomposition of kernel to do simulation.
- Main challange is to find feasible/tractable kernels for DPP models.
- We have two good choices so far:
  - A spectral model (incl. most repulsive DPP).
  - Multiquadric model.
- The multiquadric model has:

$$C_0 = \rho R_0, \quad R_0(s) = \frac{(1-\delta)^{2\tau}}{(1+\delta^2 - 2\delta\cos s)^{\tau}}, \quad \text{for } 0 < \rho \leq \rho_{\max}(R_0),\ \tau > 0,\ 0 < \delta < 1.$$

$$K_{mq}(s) = K_{Pois}(s) - 2\pi \frac{(1+\delta)(1-\delta)}{2\delta(1-2\tau)}\left(\left(\frac{1+\delta^2-2\delta}{1+\delta^2-2\delta\cos(s)}\right)^{2\tau-1} - 1\right), \quad \text{for } \tau \neq 1/2.$$

Plots



Northern Hemisphere of three spherical point patterns projected to the unit disc with an equal-area azimuthal projection. Each pattern is a simulated realization of a determinantal point process on the sphere with mean number of points 225.

- Left: Complete spatial randomness (Poisson process).
- Middle: Multiquadric model with $\tau = 10$ and $\delta = 0.68$.
- Right: Most repulsive DPP.

Palm distributions I

- Campbell-Mecke holds in general:

$$\mathrm{E}\sum_{\mathbf{x}\in\mathbf{X}} h(\mathbf{x},\mathbf{X}\setminus\{\mathbf{x}\}) = \int \mathrm{E}h(\mathbf{x},\mathbf{X}^!_{\mathbf{x}})\rho(\mathbf{x})\,\mathrm{d}\nu(\mathbf{x})$$

- As we have learned from Rasmus:

$$f^!_{\mathbf{x}}(\{\mathbf{x}_1,\dots,\mathbf{x}_n\}) = f(\{\mathbf{x},\mathbf{x}_1,\dots,\mathbf{x}_n\})/\rho(\mathbf{x}).$$

- Homogeneous (i.e. isotropic) processes in the remainder.
- Palm distribution rotation invariant (R is an arbitrary rotation):

$$\begin{aligned} f^!_{R\mathbf{x}}(\{\mathbf{x}_1,\dots,\mathbf{x}_n\}) &= f(\{\mathbf{x}_1,\dots,\mathbf{x}_n,R\mathbf{x}\})/\rho \\ &= f(\{R^\top\mathbf{x}_1,\dots,R^\top\mathbf{x}_n,R^\top R\mathbf{x}\})/\rho \\ &= f^!_{\mathbf{x}}(R^\top\{\mathbf{x}_1,\dots,\mathbf{x}_n\}) \end{aligned}$$

- Palm distribution at north pole $\mathbf{e}$ ($R^\top_{\mathbf{x}}$ rotates $\mathbf{x}$ to $\mathbf{e}$):

$$\mathrm{P}\left(\mathbf{X}^!_{\mathbf{e}}\in F\right) = \frac{1}{\nu(A)\rho}\mathrm{E}\sum_{\mathbf{x}\in\mathbf{X}\cap A} 1\left[R^\top_{\mathbf{x}}(\mathbf{X}\setminus\{\mathbf{x}\})\in F\right]$$

Palm distributions II

- **Proposition 1.** Suppose $\mathbf{X}$ is isotropic with intensity $\rho > 0$ and its distribution is absolutely continuous with respect to the unit rate Poisson process on $\mathbb{S}^2$. Then for ν almost all $\mathbf{x} \in \mathbb{S}^2$, $\mathbf{X}^!_{\mathbf{x}}$ is distributed as $R_{\mathbf{x}} \mathbf{X}^!_{\mathbf{e}}$, where the distribution of $\mathbf{X}^!_{\mathbf{e}}$ is given above. So if k is a non-negative measurable function, then for ν almost all $\mathbf{x} \in \mathbb{S}^2$,

$$\mathrm{E}k\left(R_{\mathbf{x}}^{\top}\mathbf{X}^!_{\mathbf{x}}\right) = \mathrm{E}k\left(\mathbf{X}^!_{\mathrm{e}}\right).$$

K-function

We define the *K-function* by

$$K(t) = \frac{1}{\rho} \mathrm{E} \sum_{\mathbf{x} \in \mathbf{X}_{\mathbf{e}}^!} 1\left[s(\mathbf{e}, \mathbf{x}) \leq t\right] = \frac{1}{\rho} \mathrm{E} \sum_{\mathbf{y} \in \mathbf{X}_{\mathbf{x}}^!} 1[s(\mathbf{x}, \mathbf{y}) \leq t], \quad \text{for } 0 \leq t \leq \pi, \ \mathbf{x} \in \mathbb{S}^2.$$

Then

$$\rho^2 \nu(A) K(t) = \mathrm{E} \sum_{\mathbf{x} \in \mathbf{X} \cap A} \sum_{\mathbf{y} \in \mathbf{X} \setminus \{\mathbf{x}\}} 1\left[s(\mathbf{x}, \mathbf{y}) \leq t\right]$$

For $A \subseteq \mathbb{S}^2$ we use the estimator:

$$\widehat{K}(t) = \frac{\nu(A)}{N(A)(N(A)-1)} \sum_{\substack{\mathbf{x} \in \mathbf{X} \cap A \\ \mathbf{y} \in \mathbf{X} \cap A_{\ominus t}}}^{\neq} 1\left[s(\mathbf{x}, \mathbf{y}) \leq t\right]$$

provided $N(A) > 1$.
Corresponds to estimating ρ^2 by $N(A)(N(A)-1)/\nu(A)^2$, which is unbiased for Poisson process.
The corresponding estimator when $A = \mathbb{S}^2$ and no edge correction is needed was proposed by Robeson et al. (2014).
They also noted that:

$$K(t) = 2\pi \int_0^t g_0(s) \sin s \, \mathrm{d}s, \quad \text{for } 0 \leq t \leq \pi,$$

and $K_{\mathrm{Pois}}(t) = 2\pi(1 - \cos t)$.

Normalization of K-function

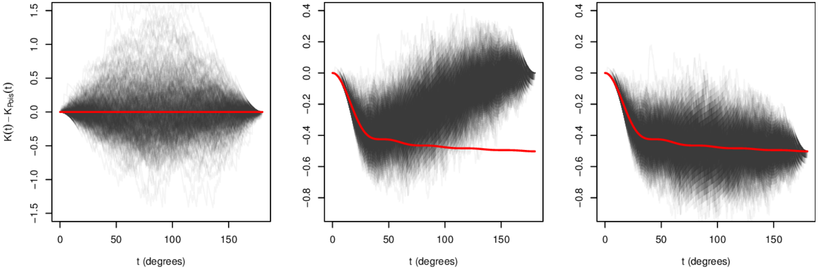
At a first glance the normalization seems unproblematic since:

$$\widehat{K}(\pi) = K_{\mathrm{Pois}}(\pi) = 4\pi$$

However in general

$$K(\pi) = 4\pi + \frac{1}{\rho}\left(\frac{\mathsf{Var}(N)}{\mathrm{E}(N)} - 1\right)$$

For a DPP we observed the following



- Left: Poisson model and usual $\widehat{K}$.
- Middle: Most repulsive DPP and usual $\widehat{K}$.
- Right: Most repulsive DPP and modified estimator ($\widehat{\rho^2} = \frac{N(A)^2}{\nu(A)^2}$)

Software

Why???

- Most important: Because large scale global phenomena requires us to use great circle distances.
- Also nice: To make your lives easier with all the GPS type data.

R package `spatstatSphere` is very much a work in progress right now!