Determinantal point processes

statistical modeling and inference

November 27, 2014

Jesper Møller
jm@math.aau.dk

Department of Mathematical Sciences
Aalborg University
Denmark

AALBORG UNIVERSITY
DENMARK

Joint work

Determinantal point processes on $\mathbb{R}^d$:

F. Lavancier, J. Møller and E. Rubak (2015). Determinantal point process models and statistical inference. To appear in *Journal of Royal Statistical Society: Series B (Statistical Methodology)*.

F. Lavancier, J. Møller and E. Rubak (2014). Determinantal point process models and statistical inference: Extended version (61 pages). Available at arXiv:1205.4818.

Joint work

Determinantal point processes on $\mathbb{R}^d$:

F. Lavancier, J. Møller and E. Rubak (2015). Determinantal point process models and statistical inference. To appear in *Journal of Royal Statistical Society: Series B (Statistical Methodology)*.

F. Lavancier, J. Møller and E. Rubak (2014). Determinantal point process models and statistical inference: Extended version (61 pages). Available at arXiv:1205.4818.

Determinantal point processes on $\mathbb{S}^d$ (in progress):

Collaborators: Emilio Porcu, University Federico Santa Maria, Valparaiso (Chile), Morten Nielsen and Ege Rubak, Dept. of Mathematical Sciences, Aalborg University.

Agenda

Definition, existence and basic properties

Stationary DPPs and approximations

Parametric models

Simulation

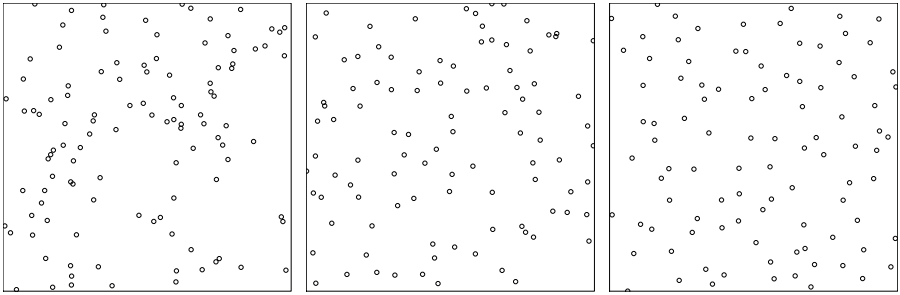
Stationary data example

Non-stationary data example

DPPs on the sphere (on going research project)

Concluding remarks

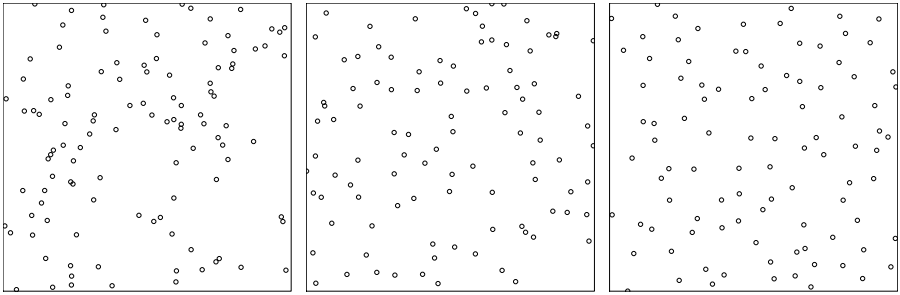
Introduction



Poisson DPP DPP with stronger inhibition

- Determinantal point processes (DPP) are inhibitive/regular/repulsive point processes.

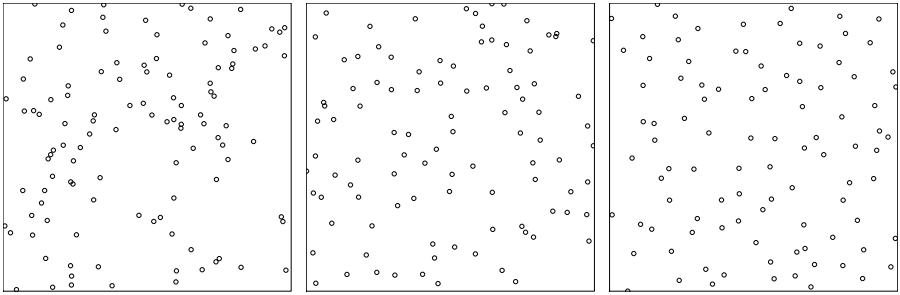
Introduction



Poisson DPP DPP with stronger inhibition

- Determinantal point processes (DPP) are inhibitive/regular/repulsive point processes.
- Introduced by O. Macchi in 1975 to model fermions in quantum mechanics.

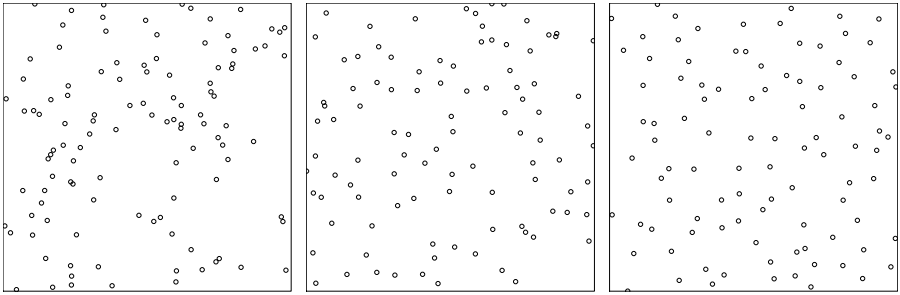
Introduction



Poisson | DPP | DPP with stronger inhibition

- Determinantal point processes (DPP) are inhibitive/regular/repulsive point processes.
- Introduced by O. Macchi in 1975 to model fermions in quantum mechanics.
- Several theoretical studies appeared in the 2000's.

Introduction



Poisson | DPP | DPP with stronger inhibition

- Determinantal point processes (DPP) are inhibitive/regular/repulsive point processes.
- Introduced by O. Macchi in 1975 to model fermions in quantum mechanics.
- Several theoretical studies appeared in the 2000's.
- **Statistical models and inference have so far been largely unexplored.**

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\ldots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}_{\{\|x_i-x_j\|\le R\}}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\ldots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \ldots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\dots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}\{\|x_i-x_j\|\le R\}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\dots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \dots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

- The normalizing constant $c(r, R, \beta, \gamma)$ is intractable.

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\ldots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}_{\{\|x_i-x_j\|\le R\}}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\ldots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \ldots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

- The normalizing constant $c(r, R, \beta, \gamma)$ is intractable.
- Interpretations? We don't know the intensity or any other moment properties.

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\ldots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}\{\|x_i-x_j\|\le R\}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\ldots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \ldots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

- The normalizing constant $c(r, R, \beta, \gamma)$ is intractable.
- Interpretations? We don't know the intensity or any other moment properties.
- We don't know the distribution of the number of points.

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\ldots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}\{\|x_i-x_j\|\le R\}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\ldots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \ldots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

- The normalizing constant $c(r, R, \beta, \gamma)$ is intractable.
- Interpretations? We don't know the intensity or any other moment properties.
- We don't know the distribution of the number of points.
- (Long) MCMC runs are needed...

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\ldots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}\{\|x_i-x_j\|\le R\}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\ldots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \ldots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

- The normalizing constant $c(r, R, \beta, \gamma)$ is intractable.
- Interpretations? We don't know the intensity or any other moment properties.
- We don't know the distribution of the number of points.
- (Long) MCMC runs are needed...
- We have ignored edge effects: the restriction to $B \subset S$ ($B \neq S$) is not a Strauss hard-core process.

Gibbs point processes — the usual class of point processes used for modelling inhibition

Example: *Strauss hard-core process*

$$f(\{x_1,\ldots,x_n\}) = \frac{1}{c(r,R,\beta,\gamma)}\beta^n \prod_{i<j} \gamma^{\mathbf{1}\{\|x_i-x_j\|\le R\}} \mathbf{1}_{\{\|x_i-x_j\|>r\}}, \qquad \{x_1,\ldots,x_n\} \subset S,$$

where $S \subset \mathbb{R}^d$ is compact; $n = 0, 1, \ldots$; $0 \le r < R$, $\beta > 0$, $0 \le \gamma \le 1$ are parameters; the density is w.r.t. the unit rate Poisson process.

- The normalizing constant $c(r, R, \beta, \gamma)$ is intractable.
- Interpretations? We don't know the intensity or any other moment properties.
- We don't know the distribution of the number of points.
- (Long) MCMC runs are needed...
- We have ignored edge effects: the restriction to $B \subset S$ ($B \neq S$) is not a Strauss hard-core process.
- On $\mathbb{R}^d$ a 'local specification' is needed and the issue of phase transition has to be clarified.

Notation

- X : spatial point process on $\mathbb{R}^d$

Notation

- X : spatial point process on $\mathbb{R}^d$
- For any Borel set $B \subseteq \mathbb{R}^d$, $X_B = X \cap B$.

Notation

- X : spatial point process on $\mathbb{R}^d$
- For any Borel set $B \subseteq \mathbb{R}^d$, $X_B = X \cap B$.
- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order joint intensity* of X:

$$\mathrm{E}\left[\#X_{B_1} \cdots \#X_{B_n}\right] = \int_{B_1} \cdots \int_{B_n} \rho^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

  for disjoint Borel sets $B_1, \ldots, B_n \subseteq \mathbb{R}^d$.

  - Intuitively,

$$\rho^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

    is the probability that for each $i = 1, \ldots, n$,
    X has a point in a region around x_i of volume $\mathrm{d}x_i$.

Notation

- X : spatial point process on $\mathbb{R}^d$
- For any Borel set $B \subseteq \mathbb{R}^d$, $X_B = X \cap B$.
- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order joint intensity* of X:

$$\mathrm{E}\left[\#X_{B_1} \cdots \#X_{B_n}\right] = \int_{B_1} \cdots \int_{B_n} \rho^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$

  for disjoint Borel sets $B_1, \ldots, B_n \subseteq \mathbb{R}^d$.

  - Intuitively,
    $$\rho^{(n)}(x_1, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n$$
    is the probability that for each $i = 1, \ldots, n$,
    X has a point in a region around x_i of volume $\mathrm{d}x_i$.
  - In particular $\rho = \rho^{(1)}$ is the *intensity function*.

Definition

Let C be a function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if

$$\rho^{(n)}(x_1, \ldots, x_n) = \det\{C(x_i, x_j\}_{i,j=1,\ldots,n}, \quad n = 1, 2, \ldots$$

Definition

Let C be a function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if

$$\rho^{(n)}(x_1, \ldots, x_n) = \det\{C(x_i, x_j\}_{i,j=1,\ldots,n}, \quad n = 1, 2, \ldots$$

- The Poisson process with intensity $\rho(x)$ is the special case where $C(x, x) = \rho(x)$ and $C(x, y) = 0$ if $x \neq y$.

Definition

Let C be a function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if

$$\rho^{(n)}(x_1, \ldots, x_n) = \det\{C(x_i, x_j\}_{i,j=1,\ldots,n}, \quad n = 1, 2, \ldots$$

- The Poisson process with intensity $\rho(x)$ is the special case where $C(x, x) = \rho(x)$ and $C(x, y) = 0$ if $x \neq y$.
- For existence, conditions on the kernel C are mandatory.
  E.g. $\det\{C(x_i, x_j\}_{i,j=1,\ldots,n} \geq 0$.

Jesper Møller

Dept. of Mathematical Sciences
Aalborg University
Denmark

Definition of a DPP

Definition

Let C be a function $\mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
X is a *determinantal point process* with *kernel* C, denoted $X \sim \mathrm{DPP}(C)$, if

$$\rho^{(n)}(x_1, \ldots, x_n) = \det\{C(x_i, x_j\}_{i,j=1,\ldots,n}, \quad n = 1, 2, \ldots$$

- The Poisson process with intensity $\rho(x)$ is the special case where $C(x, x) = \rho(x)$ and $C(x, y) = 0$ if $x \neq y$.
- For existence, conditions on the kernel C are mandatory.
  E.g. $\det\{C(x_i, x_j\}_{i,j=1,\ldots,n} \geq 0$.

For ease of exposition assume

(C1) C is a continuous (complex) covariance function.

Basic properties (if $X \sim DPP(C)$ exists)

- The intensity of X is $\rho(x) = C(x, x)$.

Basic properties (if $X \sim DPP(C)$ exists)

- The intensity of X is $\rho(x) = C(x, x)$.
- *Inhibition*, since
$$\rho^{(n)}(x_1, \ldots, x_n) \leq \rho(x_1) \cdots \rho(x_n)$$
with equality iff X is a Poisson process with intensity function ρ.

Basic properties (if $X \sim DPP(C)$ exists)

- The intensity of X is $\rho(x) = C(x, x)$.
- *Inhibition*, since
$$\rho^{(n)}(x_1, \ldots, x_n) \leq \rho(x_1) \cdots \rho(x_n)$$
with equality iff X is a Poisson process with intensity function ρ.
- The *pair correlation function* is
$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)} = 1 - |R(x, y)|^2 \leq 1$$
where R is the correlation function corresponding to C.

Basic properties (if $X \sim DPP(C)$ exists)

- The intensity of X is $\rho(x) = C(x, x)$.
- *Inhibition*, since
$$\rho^{(n)}(x_1, \ldots, x_n) \leq \rho(x_1) \cdots \rho(x_n)$$
with equality iff X is a Poisson process with intensity function ρ.
- The *pair correlation function* is
$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)} = 1 - |R(x, y)|^2 \leq 1$$
where R is the correlation function corresponding to C.
- Any *smooth transformation* or *independent thinning* of X is still a DPP with an explicitly given kernel.

Basic properties (if $X \sim DPP(C)$ exists)

- The intensity of X is $\rho(x) = C(x, x)$.
- *Inhibition*, since
$$\rho^{(n)}(x_1, \ldots, x_n) \leq \rho(x_1) \cdots \rho(x_n)$$
with equality iff X is a Poisson process with intensity function ρ.
- The *pair correlation function* is
$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)} = 1 - |R(x, y)|^2 \leq 1$$
where R is the correlation function corresponding to C.
- Any *smooth transformation* or *independent thinning* of X is still a DPP with an explicitly given kernel.
- The *restriction* to any Borel set $B \subset \mathbb{S}^d$ is a DPP with kernel $C_B(x, y) = C(x, y)$ if $x, y \in B$ and $C_B(x, y) = 0$ else.

Basic properties (if $X \sim DPP(C)$ exists)

- The intensity of X is $\rho(x) = C(x, x)$.
- *Inhibition*, since
$$\rho^{(n)}(x_1, \ldots, x_n) \leq \rho(x_1) \cdots \rho(x_n)$$
with equality iff X is a Poisson process with intensity function ρ.
- The *pair correlation function* is
$$g(x, y) := \frac{\rho^{(2)}(x, y)}{\rho(x)\rho(y)} = 1 - \frac{C(x, y)C(y, x)}{C(x, x)C(y, y)} = 1 - |R(x, y)|^2 \leq 1$$
where R is the correlation function corresponding to C.
- Any *smooth transformation* or *independent thinning* of X is still a DPP with an explicitly given kernel.
- The *restriction* to any Borel set $B \subset \mathbb{S}^d$ is a DPP with kernel $C_B(x, y) = C(x, y)$ if $x, y \in B$ and $C_B(x, y) = 0$ else.
- Given a kernel C, there *exists at most one* $\mathrm{DPP}(C)$.

Existence

By Mercer's theorem, for any compact set $S \subset \mathbb{R}^d$, C restricted to $S \times S$, denoted C_S, has a *spectral representation*,

Existence

By Mercer's theorem, for any compact set $S \subset \mathbb{R}^d$, C restricted to $S \times S$, denoted C_S, has a *spectral representation*,

$$C_S(x,y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S,$$

where $\lambda_k^S \geq 0$ and $\{\phi_k^S\}$ is a set of orthonormal basis functions for $L^2(S)$, i.e.,

$$\int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, \mathrm{d}x = \mathbf{1}_{\{k=l\}}.$$

Existence

By Mercer's theorem, for any compact set $S \subset \mathbb{R}^d$, C restricted to $S \times S$, denoted C_S, has a *spectral representation*,

$$C_S(x,y) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S,$$

where $\lambda_k^S \geq 0$ and $\{\phi_k^S\}$ is a set of orthonormal basis functions for $L^2(S)$, i.e.,

$$\int_S \phi_k^S(x) \overline{\phi_l^S(x)} \, \mathrm{d}x = \mathbf{1}_{\{k=l\}}.$$

Theorem (Macchi, 1975)

Under (C1), existence of $DPP(C)$ is equivalent to :

$$\lambda_k^S \leq 1 \text{ for all compact } S \subset \mathbb{R}^d \text{ and all } k.$$

Density on a compact set S

Let $X \sim \mathrm{DPP}(C)$ and $S \subset \mathbb{R}^d$ be any compact set.

Density on a compact set S

Let $X \sim \mathrm{DPP}(C)$ and $S \subset \mathbb{R}^d$ be any compact set.

Theorem (Macchi (1975))

If $\lambda_k^S < 1\ \forall k$, then $X_S \ll \mathit{Poisson}(S, 1)$, with density

$$f(\{x_1, \ldots, x_n\}) = \exp(|S| - D) \det\{C(x_i, x_j)\}_{i,j=1,\ldots,n},$$

where $D = -\sum_{k=1}^{\infty} \log(1 - \lambda_k^S)$ and $\tilde{C} : S \times S \to \mathbb{C}$ is given by

$$\tilde{C}(x, y) = \sum_{k=1}^{\infty} \tilde{\lambda}_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \qquad \tilde{\lambda}_k^S = \frac{\lambda_k^S}{1 - \lambda_k^S}.$$

Density on a compact set S

Let $X \sim \mathrm{DPP}(C)$ and $S \subset \mathbb{R}^d$ be any compact set.

Theorem (Macchi (1975))

If $\lambda_k^S < 1\ \forall k$, then $X_S \ll \mathit{Poisson}(S, 1)$, with density

$$f(\{x_1, \ldots, x_n\}) = \exp(|S| - D) \det\{C(x_i, x_j)\}_{i,j=1,\ldots,n},$$

where $D = -\sum_{k=1}^{\infty} \log(1 - \lambda_k^S)$ and $\tilde{C} : S \times S \to \mathbb{C}$ is given by

$$\tilde{C}(x, y) = \sum_{k=1}^{\infty} \tilde{\lambda}_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \qquad \tilde{\lambda}_k^S = \frac{\lambda_k^S}{1 - \lambda_k^S}.$$

- Thus to calculate the density/likelihood we need the spectral representation.

Density on a compact set S

Let $X \sim \mathrm{DPP}(C)$ and $S \subset \mathbb{R}^d$ be any compact set.

Theorem (Macchi (1975))

If $\lambda_k^S < 1\ \forall k$, then $X_S \ll \mathit{Poisson}(S, 1)$, with density

$$f(\{x_1, \ldots, x_n\}) = \exp(|S| - D) \det\{C(x_i, x_j)\}_{i,j=1,\ldots,n},$$

where $D = -\sum_{k=1}^{\infty} \log(1 - \lambda_k^S)$ and $\tilde{C} : S \times S \to \mathbb{C}$ is given by

$$\tilde{C}(x, y) = \sum_{k=1}^{\infty} \tilde{\lambda}_k^S \phi_k^S(x) \overline{\phi_k^S(y)}, \qquad \tilde{\lambda}_k^S = \frac{\lambda_k^S}{1 - \lambda_k^S}.$$

- Thus to calculate the density/likelihood we need the spectral representation.
- Conversely, existence of X_S is ensured by that

$$\lambda_k^S = \frac{\tilde{\lambda}_k^S}{1 + \tilde{\lambda}_k^S} < 1.$$

Simulation

Let $X \sim \mathrm{DPP}(C)$. We want to simulate X_S for $S \subset \mathbb{R}^d$ compact.

Simulation

Let $X \sim \mathrm{DPP}(C)$. We want to simulate X_S for $S \subset \mathbb{R}^d$ compact.

Theorem (Hough et al. (2006))

Let $B_1, B_2, \ldots$ be independent Bernoulli variables with means $\lambda_1^S, \lambda_1^S, \ldots$, and

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Simulation

Let $X \sim \mathrm{DPP}(C)$. We want to simulate X_S for $S \subset \mathbb{R}^d$ compact.

Theorem (Hough et al. (2006))

Let $B_1, B_2, \ldots$ be independent Bernoulli variables with means $\lambda_1^S, \lambda_1^S, \ldots$, and

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Then $DPP(C_S) \stackrel{d}{=} DPP(K)$.

Simulation

Let $X \sim \mathrm{DPP}(C)$. We want to simulate X_S for $S \subset \mathbb{R}^d$ compact.

Theorem (Hough et al. (2006))

Let $B_1, B_2, \ldots$ be independent Bernoulli variables with means $\lambda_1^S, \lambda_1^S, \ldots$, and

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Then $DPP(C_S) \stackrel{d}{=} DPP(K)$.

The algorithm starts by producing n points:

$$n \sim \sum_{k=1}^{\infty} B_k, \quad \mathrm{E}[n] = \sum_{k=1}^{\infty} \lambda_k^S, \quad \mathrm{Var}[n] = \sum_{k=1}^{\infty} \lambda_k^S (1 - \lambda_k^S).$$

Let $X \sim \mathrm{DPP}(C)$. We want to simulate X_S for $S \subset \mathbb{R}^d$ compact.

Theorem (Hough et al. (2006))

Let $B_1, B_2, \ldots$ be independent Bernoulli variables with means $\lambda_1^S, \lambda_1^S, \ldots$, and

$$K(x,y) = \sum_{k=1}^{\infty} B_k \phi_k^S(x) \overline{\phi_k^S(y)}, \quad (x,y) \in S \times S.$$

Then $DPP(C_S) \stackrel{d}{=} DPP(K)$.

The algorithm starts by producing n points:

$$n \sim \sum_{k=1}^{\infty} B_k, \quad \mathrm{E}[n] = \sum_{k=1}^{\infty} \lambda_k^S, \quad \mathrm{Var}[n] = \sum_{k=1}^{\infty} \lambda_k^S (1 - \lambda_k^S).$$

NB: Since C is continuous,

$$\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty.$$

Simulation (cont'd)

Effectively we pick out $n < \infty$ eigenfunctions with probability according to their eigenvalues and simulate the DPP with finite rank kernel

$$K(x,y) = \sum_{k:\, B_k=1} \phi_k^S(x)\overline{\phi_k^S(y)} = \sum_{i=1}^{n} \phi_{k_i}^S(x)\overline{\phi_{k_i}^S(y)}, \quad (x,y) \in S \times S.$$

Simulation (cont'd)

Effectively we pick out $n < \infty$ eigenfunctions with probability according to their eigenvalues and simulate the DPP with finite rank kernel

$$K(x,y) = \sum_{k:\, B_k=1} \phi_k^S(x)\overline{\phi_k^S(y)} = \sum_{i=1}^{n} \phi_{k_i}^S(x)\overline{\phi_{k_i}^S(y)}, \quad (x,y) \in S \times S.$$

This is a *projection kernel*, and the corresponding DPP can be simulated: The algorithm basically consists of taking a quite abstract procedure described by Hough et al. (2006) and translating it into implementable linear algebra.

Simulation (cont'd)

Effectively we pick out $n < \infty$ eigenfunctions with probability according to their eigenvalues and simulate the DPP with finite rank kernel

$$K(x,y) = \sum_{k:\, B_k=1} \phi_k^S(x)\overline{\phi_k^S(y)} = \sum_{i=1}^{n} \phi_{k_i}^S(x)\overline{\phi_{k_i}^S(y)}, \quad (x,y) \in S \times S.$$

This is a *projection kernel*, and the corresponding DPP can be simulated: The algorithm basically consists of taking a quite abstract procedure described by Hough et al. (2006) and translating it into implementable linear algebra.

This leads to simulation of the first point, the second given the first point, the third given the first and second points,...

At each step we have been using rejection sampling...

Status

- A DPP on $\mathbb{R}^d$ is specified through a continuous (complex) covariance function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.

Status

- A DPP on $\mathbb{R}^d$ is specified through a continuous (complex) covariance function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
- C determines the moment properties of the DPP.

Status

- A DPP on $\mathbb{R}^d$ is specified through a continuous (complex) covariance function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
- C determines the moment properties of the DPP.
- Given the spectral representation of C on a compact set S we
  - have a simple existence condition,
  - know the distribution of the number of points falling in S,
  - can simulate the process on S,
  - can calculate the density/likelihood.

Status

- A DPP on $\mathbb{R}^d$ is specified through a continuous (complex) covariance function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$.
- C determines the moment properties of the DPP.
- Given the spectral representation of C on a compact set S we
  - have a simple existence condition,
  - know the distribution of the number of points falling in S,
  - can simulate the process on S,
  - can calculate the density/likelihood.

Typically we don't know the spectral representation!

Stationary kernels

Consider a stationary kernel: $C(x, y) = C_0(x - y), \quad x, y \in \mathbb{R}^d$.

Stationary kernels

Consider a stationary kernel: $C(x, y) = C_0(x - y), \quad x, y \in \mathbb{R}^d$.

Its Fourier transform (or spectral density) is:

$$\varphi(x) = \int C_0(t) \mathrm{e}^{-2\pi \mathrm{i} x \cdot t} \, \mathrm{d}t, \quad x \in \mathbb{R}^d.$$

Stationary kernels

Consider a stationary kernel: $C(x, y) = C_0(x - y), \quad x, y \in \mathbb{R}^d$.

Its Fourier transform (or spectral density) is:

$$\varphi(x) = \int C_0(t) \mathrm{e}^{-2\pi \mathrm{i} x \cdot t} \, \mathrm{d}t, \quad x \in \mathbb{R}^d.$$

Theorem

Under (C1), if $C_0 \in L^2(\mathbb{R}^d)$, then existence of $DPP(C_0)$ is equivalent to

$$\varphi \leq 1.$$

Stationary kernels

Consider a stationary kernel: $C(x, y) = C_0(x - y), \quad x, y \in \mathbb{R}^d$.

Its Fourier transform (or spectral density) is:

$$\varphi(x) = \int C_0(t) \mathrm{e}^{-2\pi \mathrm{i} x \cdot t} \, \mathrm{d}t, \quad x \in \mathbb{R}^d.$$

Theorem

Under (C1), if $C_0 \in L^2(\mathbb{R}^d)$, then existence of $DPP(C_0)$ is equivalent to

$$\varphi \leq 1.$$

$\rightarrow$ *This induces a restriction on the parameter space*:

That is, there is a trade-off between strong inhibiton and large intensity.

In practice, this restriction implies that if the intensity is large the range (effective support) of C_0 must be small.

Approximation

WLOG consider $S = [-1/2, 1/2]^d$.

Approximation

WLOG consider $S = [-1/2, 1/2]^d$. Approximate X_S by $X^{\text{app}} \sim DPP_S(C_{\text{app}})$ where

$$C_{\text{app}}(x, y) = \sum_{k \in \mathbb{Z}^d} \varphi(k) \mathrm{e}^{2\pi \mathrm{i} k \cdot (x-y)}, \quad x, y \in S.$$

Approximation

WLOG consider $S = [-1/2, 1/2]^d$. Approximate X_S by $X^{\text{app}} \sim DPP_S(C_{\text{app}})$ where

$$C_{\text{app}}(x, y) = \sum_{k \in \mathbb{Z}^d} \varphi(k) \mathrm{e}^{2\pi \mathrm{i} k \cdot (x-y)}, \quad x, y \in S.$$

If $x - y \in S$ this is effectively the Fourier expansion

$$C(x, y) = C_0(x - y) = \sum_{k \in \mathbb{Z}^d} \alpha_k \mathrm{e}^{2\pi \mathrm{i} k \cdot (x-y)}$$

since for "most" interesting models

$$\alpha_k = \int_S C_0(t) \mathrm{e}^{-2\pi \mathrm{i} k \cdot t} \, \mathrm{d}t \approx \int_{\mathbb{R}^d} C_0(t) \mathrm{e}^{-2\pi \mathrm{i} k \cdot t} \, \mathrm{d}t = \varphi(k).$$

So we claim that $C_0(t) \approx 0$ for $t \notin S$: in practice, for any reasonable expected number of points, this is implied by the parameter restriction.

Modelling based on spectral densities

Idea: instead of modelling C_0, model λ_k^S and ϕ_k^S in

$$C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}.$$

Modelling based on spectral densities

Idea: instead of modelling C_0, model λ_k^S and ϕ_k^S in

$$C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}.$$

Following the previous approximation on the unit square:

- Choose the Fourier basis: $\phi_k^S(x) = \mathrm{e}^{-2\pi i k \cdot x}$.
- Choose $\lambda_k^S = \varphi(k)$, where φ is a spectral density with $\varphi \leq 1$.

Modelling based on spectral densities

Idea: instead of modelling C_0, model λ_k^S and ϕ_k^S in

$$C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}.$$

Following the previous approximation on the unit square:

- Choose the Fourier basis: $\phi_k^S(x) = \mathrm{e}^{-2\pi i k \cdot x}$.
- Choose $\lambda_k^S = \varphi(k)$, where φ is a spectral density with $\varphi \leq 1$.
- Then we have a well-defined DPP on S, which can easily be simulated and the density/likelihood can be evaluated exactly (up to series truncation).

Modelling based on spectral densities

Idea: instead of modelling C_0, model λ_k^S and ϕ_k^S in

$$C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}.$$

Following the previous approximation on the unit square:

- Choose the Fourier basis: $\phi_k^S(x) = \mathrm{e}^{-2\pi \mathrm{i} k \cdot x}$.
- Choose $\lambda_k^S = \varphi(k)$, where φ is a spectral density with $\varphi \leq 1$.
- Then we have a well-defined DPP on S, which can easily be simulated and the density/likelihood can be evaluated exactly (up to series truncation).
- To obtain a DPP on $\mathbb{R}^d$ start by modelling $\varphi \leq 1$.

Modelling based on spectral densities

Idea: instead of modelling C_0, model λ_k^S and ϕ_k^S in

$$C_0(y-x) = \sum_{k=1}^{\infty} \lambda_k^S \phi_k^S(x) \overline{\phi_k^S(y)}.$$

Following the previous approximation on the unit square:

- Choose the Fourier basis: $\phi_k^S(x) = \mathrm{e}^{-2\pi i k \cdot x}$.
- Choose $\lambda_k^S = \varphi(k)$, where φ is a spectral density with $\varphi \leq 1$.
- Then we have a well-defined DPP on S, which can easily be simulated and the density/likelihood can be evaluated exactly (up to series truncation).
- To obtain a DPP on $\mathbb{R}^d$ start by modelling $\varphi \leq 1$.

Main drawback:

- C_0 (and thus the moment properties) is given as an infinite sum $\rightarrow$ parameters may be harder to understand/interpret.

Intermezzo

This concludes the first part of the talk focusing on the probabilistic background and approximations for simulation and density expression.

Now we start doing statistics, so if you got lost or fell asleep you get a fresh start!

Examples of parametric models

We will focus on the following parametric models, where $\rho > 0$ is the intensity, $\alpha > 0$ is a scale/range parameter, and $\nu > 0$ is a shape parameter:

Examples of parametric models

We will focus on the following parametric models, where $\rho > 0$ is the intensity, $\alpha > 0$ is a scale/range parameter, and $\nu > 0$ is a shape parameter:

- *Whittle-Matérn model*, which includes the exponential model ($\nu = 1/2$) and the Gaussian model ($\nu = \infty$):

$$C_0(x) = \rho \frac{2^{1-\nu}}{\Gamma(\nu)} \|x/\alpha\|^\nu K_\nu(\|x/\alpha\|), \quad x \in \mathbb{R}^d,$$

  The parameter restriction is $\rho \leq \frac{\Gamma(\nu)}{\Gamma(\nu+d/2)(2\sqrt{\pi}\alpha)^d}$.

Examples of parametric models

We will focus on the following parametric models, where $\rho > 0$ is the intensity, $\alpha > 0$ is a scale/range parameter, and $\nu > 0$ is a shape parameter:

- *Whittle-Matérn model*, which includes the exponential model ($\nu = 1/2$) and the Gaussian model ($\nu = \infty$):

$$C_0(x) = \rho \frac{2^{1-\nu}}{\Gamma(\nu)} \|x/\alpha\|^\nu K_\nu(\|x/\alpha\|), \quad x \in \mathbb{R}^d,$$

  The parameter restriction is $\rho \leq \frac{\Gamma(\nu)}{\Gamma(\nu+d/2)(2\sqrt{\pi}\alpha)^d}$.

- *Power exponential spectral model*

$$\varphi(x) = \rho \frac{\Gamma(d/2+1)\alpha^d}{\pi^{d/2}\Gamma(d/\nu+1)} \exp(-\|\alpha x\|^\nu), \quad x \in \mathbb{R}^d.$$

  The parameter restriction is $\rho \leq \frac{\pi^{d/2}\Gamma(d/\nu+1)}{\Gamma(d/2)\alpha^d}$.

Parametric models in R (so far: contact Ege Rubak; later on: spatstat)

The parametric families are specified in R via the determinantal family functions (of class `detfamily`): `detGauss`, `detMatern`, `detPowerExp`. E.g:

- `model <- detGauss(rho=100, alpha=0.05, d=2)`
- `model <- detMatern(rho=100, alpha=0.03, nu=0.5, d=2)`
- `model <- detPowerExp(rho=100, alpha=0.17, nu=2, d=2)`

Parametric models in R (so far: contact Ege Rubak; later on: spatstat)

The parametric families are specified in R via the determinantal family functions (of class `detfamily`): `detGauss`, `detMatern`, `detPowerExp`. E.g:

- `model <- detGauss(rho=100, alpha=0.05, d=2)`
- `model <- detMatern(rho=100, alpha=0.03, nu=0.5, d=2)`
- `model <- detPowerExp(rho=100, alpha=0.17, nu=2, d=2)`

Extract the kernel, spectral density, pair correlation function, K-function:

- `detkernel(model)`
- `detspecden(model)`
- `pcfmodel(model)`
- `Kmodel(model)`

Simulation in R

Simply use the generic function `simulate` (then R automatically calls the function `simulate.detmodel`):

- ```
  model <- detGauss(rho=100, alpha=0.05, d=2)
  X <- simulate(model)
  ```
  ```

Simulation in R

Simply use the generic function `simulate` (then R automatically calls the function `simulate.detmodel`):

- ```
  model <- detGauss(rho=100, alpha=0.05, d=2)
  X <- simulate(model)
  ```
- Change the window (default is the unit square):
  ```
  W <- owin(poly=list(x=c(-1,0,1),y=c(0,1,0)))
  X <- simulate(model, W=W)
  ```
  ```

Simulation in R

Simply use the generic function `simulate` (then R automatically calls the function `simulate.detmodel`):

- ```
  model <- detGauss(rho=100, alpha=0.05, d=2)
  X <- simulate(model)
  ```
- Change the window (default is the unit square):
  ```
  W <- owin(poly=list(x=c(-1,0,1),y=c(0,1,0)))
  X <- simulate(model, W=W)
  ```
- Several realizations:
  ```
  X <- simulate(model, nsim=4)
  ```
  ```

Illustration of simulation algorithm

Step 1. The first point is sampled uniformly on S (stationary case).

Illustration of simulation algorithm

Step 2. The next point is sampled w.r.t. the following density:

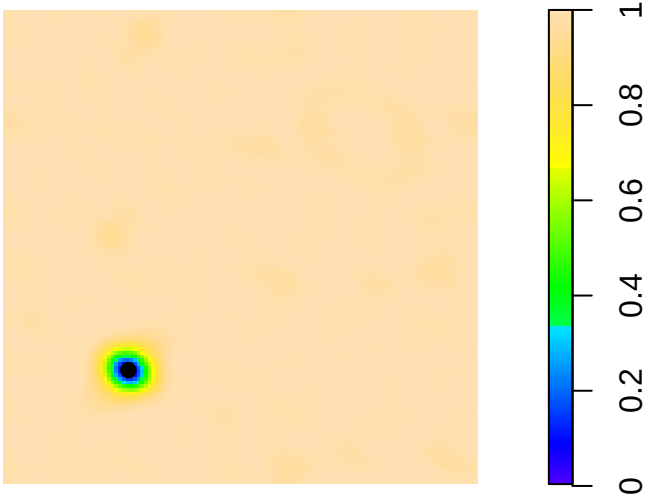
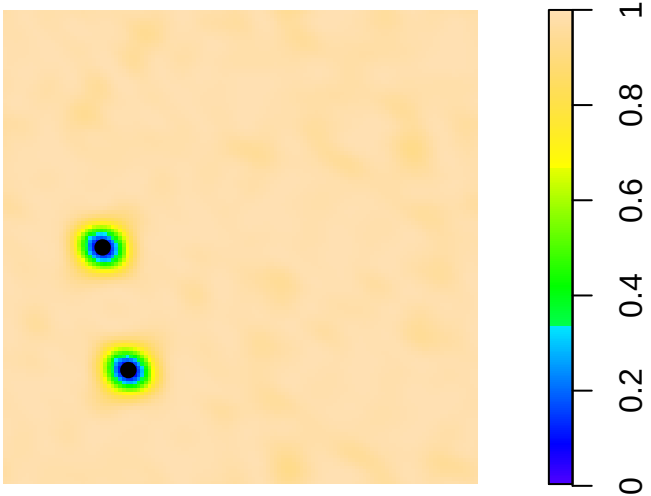


Illustration of simulation algorithm

Step 3. The next point is sampled w.r.t. the following density:



Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

Step 4. The next point is sampled w.r.t. the following density:

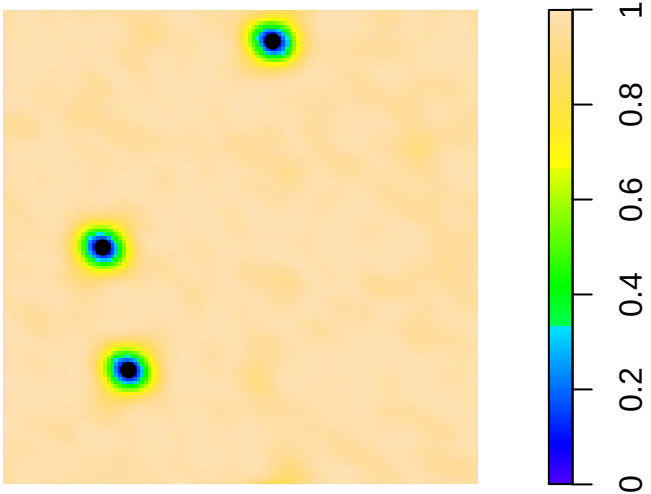


Illustration of simulation algorithm

Step 5. The next point is sampled w.r.t. the following density:

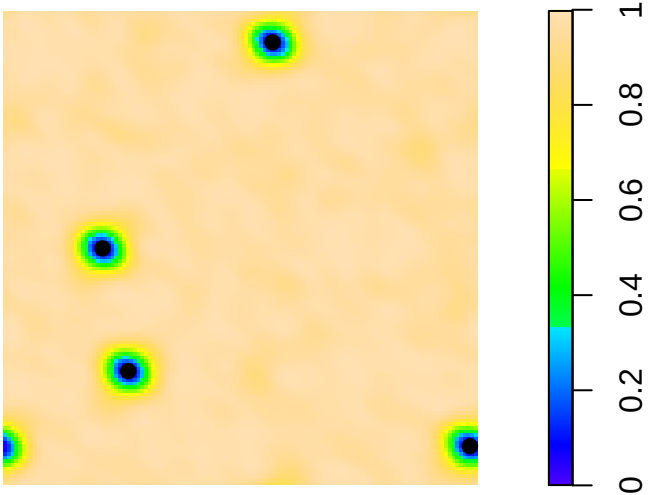
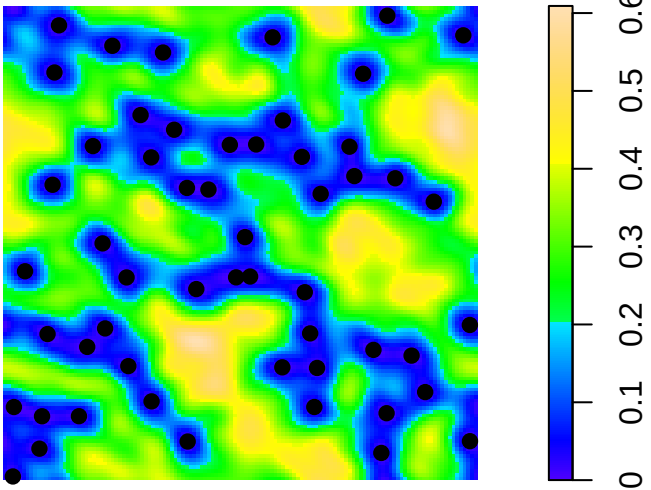


Illustration of simulation algorithm

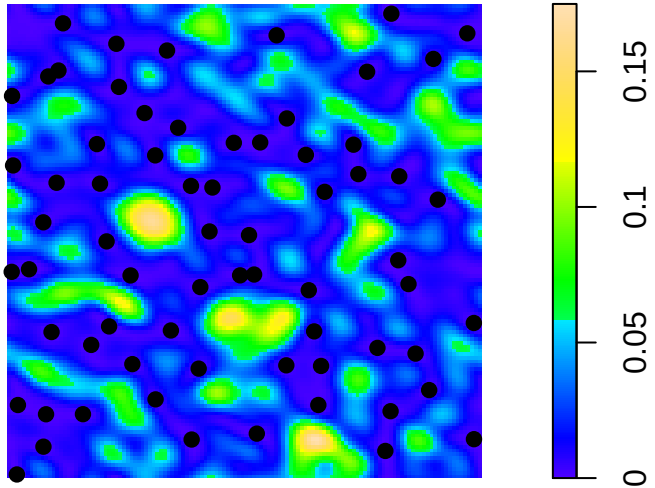
...somewhere in the middle...



Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

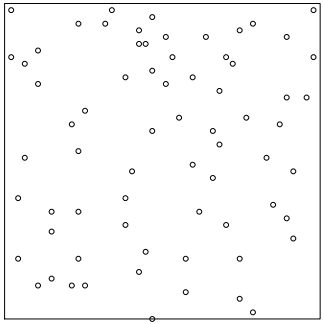
Final point is sampled w.r.t. the following density:



Spanish towns dataset

Ripley (1988): Strauss hard-core model with 4 parameters:

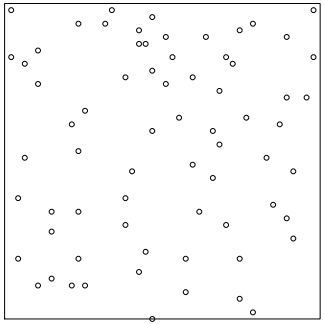
r=hard-core, R=range of interaction, β=abundance, γ=interaction.



Spanish towns dataset

Ripley (1988): Strauss hard-core model with 4 parameters:

r=hard-core, R=range of interaction, β=abundance, γ=interaction.



Following Illian et al. (2008): $\hat{r} = 0.83$, $\hat{R} = 3.5$.
Approximate likelihood method (Huang and Ogata (1999)): $\hat{\beta} = 0.12$ and $\hat{\gamma} = 0.76$.

Alternative DPP models

Gaussian, Whittle-Matérn, and power exponential spectral models fitted using the function `dppm`:

- Default estimation method is "partial likelihood" where we use $\hat{\rho} = n/|W| = 0.043$ and MLEs for the rest:
  ```
  fit <- dppm(X, detGauss())
  ```
- Full likelihood:
  ```
  fit <- dppm(X, detGauss(), method="likelihood")
  ```

Alternative DPP models

Gaussian, Whittle-Matérn, and power exponential spectral models fitted using the function `dppm`:

- Default estimation method is "partial likelihood" where we use $\hat{\rho} = n/|W| = 0.043$ and MLEs for the rest:
  ```
  fit <- dppm(X, detGauss())
  ```
- Full likelihood:
  ```
  fit <- dppm(X, detGauss(), method="likelihood")
  ```

Highest likelihood: fitted Whittle-Matérn model.

Alternative DPP models

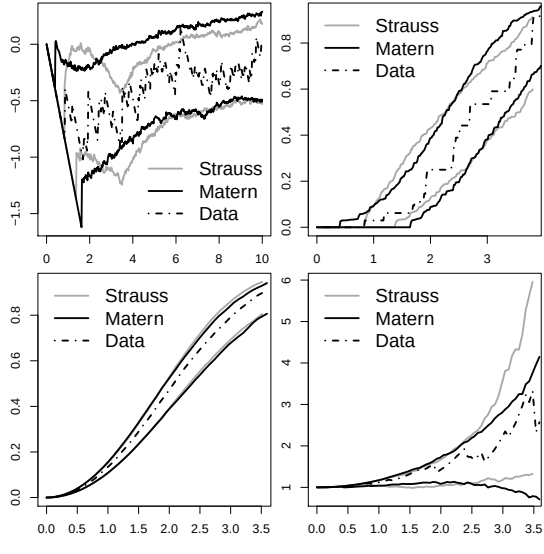
Gaussian, Whittle-Matérn, and power exponential spectral models fitted using the function `dppm`:

- Default estimation method is "partial likelihood" where we use $\hat{\rho} = n/|W| = 0.043$ and MLEs for the rest:
  ```
  fit <- dppm(X, detGauss())
  ```
- Full likelihood:
  ```
  fit <- dppm(X, detGauss(), method="likelihood")
  ```

Highest likelihood: fitted Whittle-Matérn model.

Simulation based likelihood-ratio test for the simpler Gaussian model vs the Whittle-Matérn model: $p = 3\%$.

Clockwise from top left: Non-parametric estimate of $L(r) - r$, $G(r)$, $J(r)$, $F(r)$, and simulation based 2.5% and 97.5% pointwise quantiles (based on 400 realizations).



Conclusion of data analysis

Whittle-Matérn model:

- has **less parameters**
- (arguably) provides a better fit
- has a canonical way of estimating parameters (**likelihood**)
- direct **access to the moments** (intensity, pair correlation function, ...)

Conclusion of data analysis

Whittle-Matérn model:

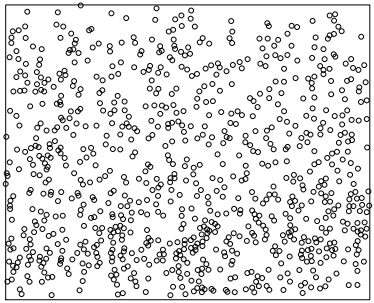
- has **less parameters**
- (arguably) provides a better fit
- has a canonical way of estimating parameters (**likelihood**)
- direct **access to the moments** (intensity, pair correlation function, ...)

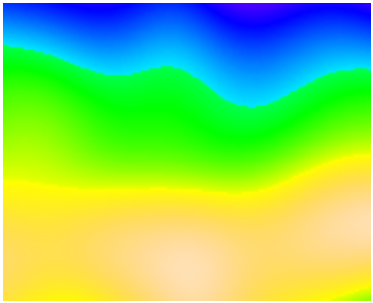
For the Strauss hard-core model

- parameter estimation relies to a certain extend on “ad-hoc” methods
- the density and moments can only be obtained by MCMC simulation.

Mucous membrane dataset

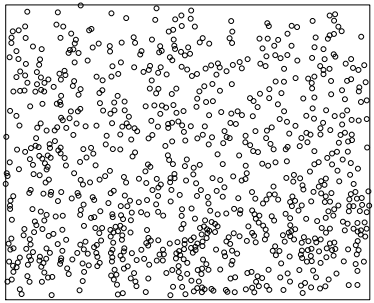
Consists of the most abundant type of cell in a bivariate point pattern analysed in Møller and Waagepetersen (2004).

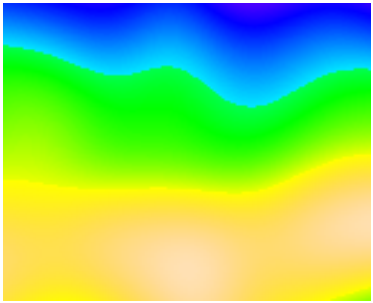




Mucous membrane dataset

Consists of the most abundant type of cell in a bivariate point pattern analysed in Møller and Waagepetersen (2004).





We use this unmarked point pattern to illustrate how an **inhomogenous DPP** can be fitted to a real dataset.

Modelling inhomogeneity

Assume *second-order intensity-reweighted stationarity* (Baddeley, Møller & Waagepetersen, 2000), i.e., the correlation function is translation invariant:

$$R(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \frac{C(x,y)}{\sqrt{\rho(x)\rho(y)}} = R_0(x-y).$$

Modelling inhomogeneity

Assume *second-order intensity-reweighted stationarity* (Baddeley, Møller & Waagepetersen, 2000), i.e., the correlation function is translation invariant:

$$R(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \frac{C(x,y)}{\sqrt{\rho(x)\rho(y)}} = R_0(x-y).$$

- Fit a parametric model to ρ depending on relevant covariates (second coordinate axis in our case).

Modelling inhomogeneity

Assume *second-order intensity-reweighted stationarity* (Baddeley, Møller & Waagepetersen, 2000), i.e., the correlation function is translation invariant:

$$R(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \frac{C(x,y)}{\sqrt{\rho(x)\rho(y)}} = R_0(x-y).$$

- Fit a parametric model to ρ depending on relevant covariates (second coordinate axis in our case).
- Use the fitted intensity to estimate the inhomogeneous g-function (or K-function).

Modelling inhomogeneity

Assume *second-order intensity-reweighted stationarity* (Baddeley, Møller & Waagepetersen, 2000), i.e., the correlation function is translation invariant:

$$R(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \frac{C(x,y)}{\sqrt{\rho(x)\rho(y)}} = R_0(x-y).$$

- Fit a parametric model to ρ depending on relevant covariates (second coordinate axis in our case).
- Use the fitted intensity to estimate the inhomogeneous g-function (or K-function).
- Fit a parametric model for R_0 via minimum contrast.

Modelling inhomogeneity

Assume *second-order intensity-reweighted stationarity* (Baddeley, Møller & Waagepetersen, 2000), i.e., the correlation function is translation invariant:

$$R(x,y) = \frac{C(x,y)}{\sqrt{C(x,x)C(y,y)}} = \frac{C(x,y)}{\sqrt{\rho(x)\rho(y)}} = R_0(x-y).$$

- Fit a parametric model to ρ depending on relevant covariates (second coordinate axis in our case).
- Use the fitted intensity to estimate the inhomogeneous g-function (or K-function).
- Fit a parametric model for R_0 via minimum contrast.
- The resulting DPP has kernel

$$\hat{C}(x,y) = \sqrt{\hat{\rho}(x)}\hat{R}_0(x-y)\sqrt{\hat{\rho}(y)}.$$

Simulation of inhomogeneous model

NB: If a 'dominating DPP' with kernel $C^{\mathrm{dom}}(x, y)$ is thinned with retention probability $\pi(x)$, the resulting process is a new DPP with kernel

$$C(x, y) = \sqrt{\pi(x)}\, C^{\mathrm{dom}}(x, y) \sqrt{\pi(y)}.$$

Simulation of inhomogeneous model

NB: If a 'dominating DPP' with kernel $C^{\mathrm{dom}}(x,y)$ is thinned with retention probability $\pi(x)$, the resulting process is a new DPP with kernel

$$C(x,y) = \sqrt{\pi(x)}C^{\mathrm{dom}}(x,y)\sqrt{\pi(y)}.$$

Thus let $\hat{\rho}_{\max} = \sup_{x\in S}\{\hat{\rho}(x)\}$ and define a stationary DPP X^{dom} with kernel

$$C^{\mathrm{dom}}(x,y) = C_0^{\mathrm{dom}}(x-y) = \hat{\rho}_{\max}\hat{R}_0(x-y).$$

Simulation of inhomogeneous model

NB: If a 'dominating DPP' with kernel $C^{\mathrm{dom}}(x,y)$ is thinned with retention probability $\pi(x)$, the resulting process is a new DPP with kernel

$$C(x,y) = \sqrt{\pi(x)}C^{\mathrm{dom}}(x,y)\sqrt{\pi(y)}.$$

Thus let $\hat{\rho}_{\max} = \sup_{x\in S}\{\hat{\rho}(x)\}$ and define a stationary DPP X^{dom} with kernel

$$C^{\mathrm{dom}}(x,y) = C_0^{\mathrm{dom}}(x-y) = \hat{\rho}_{\max}\hat{R}_0(x-y).$$

Then our fitted model is simulated by thinning X^{dom} with retention probability $\pi(x) = \hat{\rho}(x)/\hat{\rho}_{\max}$, since

$$\sqrt{\frac{\hat{\rho}(x)}{\hat{\rho}_{\max}}}C^{\mathrm{dom}}(x,y)\sqrt{\frac{\hat{\rho}(y)}{\hat{\rho}_{\max}}} = \sqrt{\hat{\rho}(x)}\hat{R}_0(x-y)\sqrt{\hat{\rho}(y)} = \hat{C}(x,y).$$

DPPs on the sphere

- On the sphere the spherical harmonics constitute a set of basis functions (given in terms of associated Legendre polynomials).

DPPs on the sphere

- On the sphere the spherical harmonics constitute a set of basis functions (given in terms of associated Legendre polynomials).
- Thus we only have to make a parametric model for the eigenvalues λ_k to have a DPP on the sphere.

DPPs on the sphere

- On the sphere the spherical harmonics constitute a set of basis functions (given in terms of associated Legendre polynomials).
- Thus we only have to make a parametric model for the eigenvalues λ_k to have a DPP on the sphere.
- There are covariance functions on the sphere with known eigenvalues. One is the Inverse MultiQuadric covariance function.

DPPs on the sphere

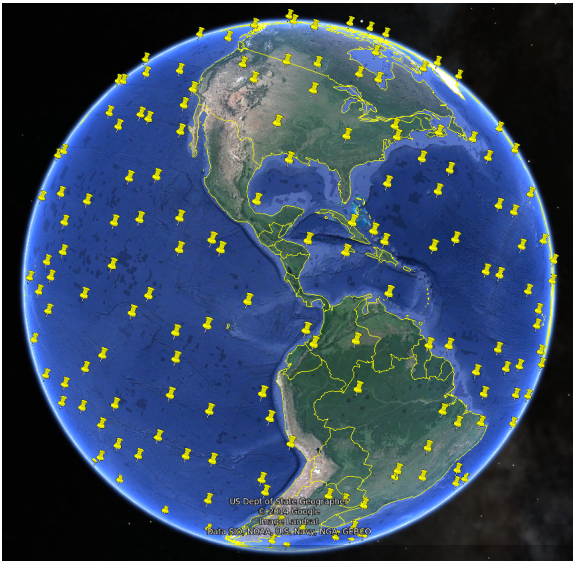
- On the sphere the spherical harmonics constitute a set of basis functions (given in terms of associated Legendre polynomials).
- Thus we only have to make a parametric model for the eigenvalues λ_k to have a DPP on the sphere.
- There are covariance functions on the sphere with known eigenvalues. One is the Inverse MultiQuadric covariance function.
- We have implemented it in R:

```
model <- detIMQ(rho=500,delta=0.998)
```

DPPs on the sphere

- On the sphere the spherical harmonics constitute a set of basis functions (given in terms of associated Legendre polynomials).
- Thus we only have to make a parametric model for the eigenvalues λ_k to have a DPP on the sphere.
- There are covariance functions on the sphere with known eigenvalues. One is the Inverse MultiQuadric covariance function.
- We have implemented it in R:
  `model <- detIMQ(rho=500,delta=0.998)`
- Simlulations rely on the previously developed code (with some modifications):
  `X <- simulate(model)`

A simulated DPP consisting of 441 points on planet Earth



DPP vs Gibbs

DPP's possess **appealing properties**:

DPP vs Gibbs

DPP's possess **appealing properties**:

- They provide flexible parametric models of repulsive point processes ('soft-core' cases and some cases with more repulsion).

DPP vs Gibbs

DPP's possess **appealing properties**:

- They provide flexible parametric models of repulsive point processes ('soft-core' cases and some cases with more repulsion).
- Easily and very quickly simulated.

DPP vs Gibbs

DPP's possess **appealing properties**:

- They provide flexible parametric models of repulsive point processes ('soft-core' cases and some cases with more repulsion).
- Easily and very quickly simulated.
- Closed form expressions for all orders of moments.

DPP vs Gibbs

DPP's possess **appealing properties**:

- They provide flexible parametric models of repulsive point processes ('soft-core' cases and some cases with more repulsion).
- Easily and very quickly simulated.
- Closed form expressions for all orders of moments.
- Closed form expression for the density of a DPP on any bounded set.

DPP vs Gibbs

DPP's possess **appealing properties**:

- They provide flexible parametric models of repulsive point processes ('soft-core' cases and some cases with more repulsion).
- Easily and very quickly simulated.
- Closed form expressions for all orders of moments.
- Closed form expression for the density of a DPP on any bounded set.
- Inference is feasible, including likelihood inference. Freely avaliable software!

DPP vs Gibbs

DPP's possess **appealing properties**:

- They provide flexible parametric models of repulsive point processes ('soft-core' cases and some cases with more repulsion).
- Easily and very quickly simulated.
- Closed form expressions for all orders of moments.
- Closed form expression for the density of a DPP on any bounded set.
- Inference is feasible, including likelihood inference. Freely avaliable software!

$\Rightarrow$ **Promising alternative to Gibbs point processes**.

Future developments

- Implementing more models (circular, generalized Cauchy, generalized sinc, Laguerre-Gauss, ...).

Future developments

- Implementing more models (circular, generalized Cauchy, generalized sinc, Laguerre-Gauss, ...).
- Implementing different algorithms for approximating the likelihood (based on FFT, convolution approximation etc).

Future developments

- Implementing more models (circular, generalized Cauchy, generalized sinc, Laguerre-Gauss, ...).
- Implementing different algorithms for approximating the likelihood (based on FFT, convolution approximation etc).
- Developing C-code for simulation and inference.

Future developments

- Implementing more models (circular, generalized Cauchy, generalized sinc, Laguerre-Gauss, ...).
- Implementing different algorithms for approximating the likelihood (based on FFT, convolution approximation etc).
- Developing C-code for simulation and inference.
- Developing and implementing more models on the sphere.

Future developments

- Implementing more models (circular, generalized Cauchy, generalized sinc, Laguerre-Gauss, ...).
- Implementing different algorithms for approximating the likelihood (based on FFT, convolution approximation etc).
- Developing C-code for simulation and inference.
- Developing and implementing more models on the sphere.
- Implementing summary statistics on the sphere.

References

[1] Hough, J. B., M. Krishnapur, Y. Peres, and B. Virág (2006).
Determinantal processes and independence.
Probability Surveys 3, 206–229.

[2] Macchi, O. (1975).
The coincidence approach to stochastic point processes.
Advances in Applied Probability 7, 83–122.

[3] McCullagh, P. and J. Møller (2006).
The permanental process.
Advances in Applied Probability 38, 873–888.

[4] Møller, J. and R. P. Waagepetersen (2004).
Statistical Inference and Simulation for Spatial Point Processes.
CRC/Chapman & Hall.

[5] Scardicchio, A., C. Zachary, and S. Torquato (2009).
Statistical properties of determinantal point processes in high-dimensional Euclidean spaces.
Physical Review E 79(4).

[6] Shirai, T. and Y. Takahashi (2003).
Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes.
Journal of Functional Analysis 2, 414–463.

[7] Soshnikov, A. (2000).
Determinantal random point fields.
Russian Math. Surveys 55, 923–975.

Thank you for you attention!

AALBORG UNIVERSITY
DENMARK