Determinantal point processes
statistical modeling and inference

August 2022

Ege Rubak
rubak@math.aau.dk

Department of Mathematical Sciences
Aalborg University
Denmark

AALBORG UNIVERSITY
DENMARK

Agenda

Definition, existence and basic properties

Stationary DPPs and approximations

Parametric models

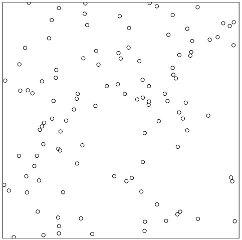
Simulation

Data example

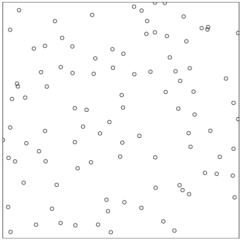
Dept. of Mathematical Sciences
Aalborg University
Denmark

Introduction

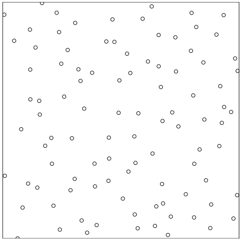
- Determinantal point processes (DPP) are inhibitive (or regular, or repulsive) point processes.
- Introduced by O. Macchi in 1975 to model fermions in quantum mechanics.
- Several theoretical studies appeared in the 2000's.
- Statistical modeling came along in the 2010's and onward.



Poisson



DPP



DPP with stronger inhibition

Ege Rubak

Dept. of Mathematical Sciences
Aalborg University
Denmark

Notation

- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order product density function* of X. 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 $\mathrm{d}x_i$.

Dept. of Mathematical Sciences
Aalborg University
Denmark

Notation

- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order product density function* of X. Intuitively,

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

  is the probability that for each $i = 1, \dots, n$,
  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*.

Notation

- For any integer $n > 0$, denote $\rho^{(n)}$ the n'th *order product density function* of X. Intuitively,

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

is the probability that for each $i = 1, \dots, n$,
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*.

- For any function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$, denote $[C](x_1, \dots, x_n)$ the $n \times n$ matrix with entries $C(x_i, x_j)$.

$$\text{Ex}: \quad [C](x_1) = C(x_1, x_1) \quad [C](x_1, x_2) = \begin{pmatrix} C(x_1, x_1) & C(x_1, x_2) \\ C(x_2, x_1) & C(x_2, x_2) \end{pmatrix}.$$

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 its product density functions satisfy

$$\rho^{(n)}(x_1, \dots, x_n) = \det[C](x_1, \dots, x_n), \quad n = 1, 2, \dots$$

For existence, conditions on the kernel C are mandatory

- C must e.g. satisfy: $\det[C](x_1, \dots, x_n) \geq 0$ for all $x_1, \dots, x_n$.
- Henceforth we assume

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

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

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

$$g(x,y) := \frac{\rho^{(2)}(x,y)}{\rho(x)\rho(y)} = 1 - \frac{C(x,y)C(y,x)}{C(x,x)C(y,y)} = 1 - |R(x,y)|^2$$

  where R is the correlation function corresponding to C.
- Thus $g \leq 1$ (i.e. inhibition) since C is Hermitian by (C1).
- If $X \sim \mathrm{DPP}(C)$, then $X_B \sim \mathrm{DPP}_B(C_B)$
- Any smooth transformation or independent thinning of a DPP is still a DPP with an explicitly given kernel.

Dept. of Mathematical Sciences
Aalborg University
Denmark

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 $\{\phi_k\}$ is a set of orthonormal basis functions on 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 $\mathrm{DPP}(C)$ is equivalent to :

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

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))

For $k \in \mathbb{N}$, let B_k be independent Bernoulli r.v. with mean λ_k^S. Define

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

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

Note that almost surely there is a finite number of ones in the Bernoulli sequence B_k since $\sum \lambda_k^S = \int_S C(x,x)\,\mathrm{d}x < \infty$.

Simulation (cont'd)

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

$$K(x, y) = \sum_{k'=1}^{n} \phi_{k'}^S(x)\overline{\phi_{k'}^S(y)}, \quad (x, y) \in S \times S.$$

This is a *projection kernel*, and the corresponding DPP can be simulated using rejection sampling.

The algorithm always produces n points. Thus,

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

Dept. of Mathematical Sciences
Aalborg University
Denmark

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 $\mathrm{DPP}(C_0)$ is equivalent to

$$\varphi \leq 1.$$

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

In practice, this restriction implies that if the intensity is large the range (effective support) of C_0 must be small. I.e. there is a trade-off between strong inhibiton and large intensity.

Approximation

Without loss of generality we consider $S = [-1/2, 1/2]^d$. To approximate X_S we consider $X^{\text{app}} \sim DPP_S(C_{\mathrm{app}})$ where

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

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 Gauss 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 $\rho \leq \frac{\Gamma(\nu)}{\Gamma(\nu+d/2)(2\sqrt{\pi}\alpha)^d}$

Dept. of Mathematical Sciences
Aalborg University
Denmark

Parametric models in R

The parametric families are specified in R via the determinantal family functions: `dppGauss`, `dppMatern`, `dppPowerExp`. E.g:

- `model <- dppGauss(lambda=100, alpha=0.05, d=2)`
- `model <- dppMatern(lambda=100, alpha=0.03, nu=0.5, d=2)`
- `model <- dppPowerExp(lambda=100, alpha=0.17, nu=2, d=2)`

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

- `dppkernel(model)`
- `dppspecden(model)`
- `pcfmodel(model)`
- `Kmodel(model)`

Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

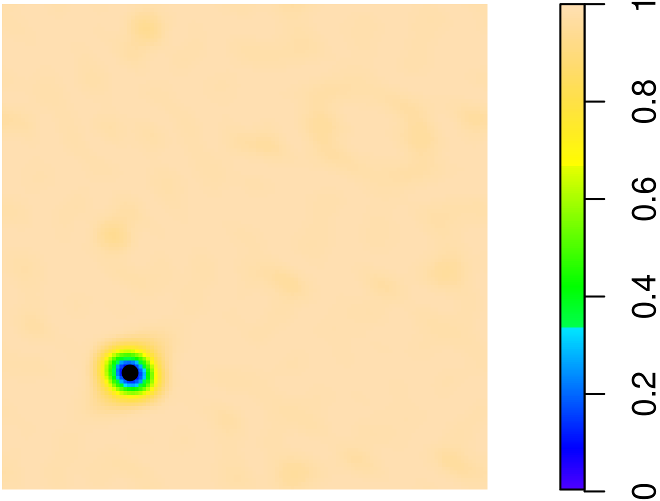
Step 1. The first point is sampled uniformly on S.

Ege Rubak

Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

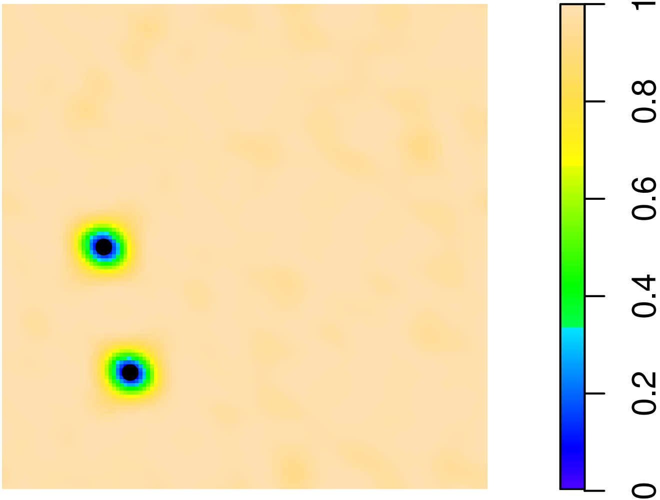
Step 2. The next point is sampled w.r.t. the following density:



Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

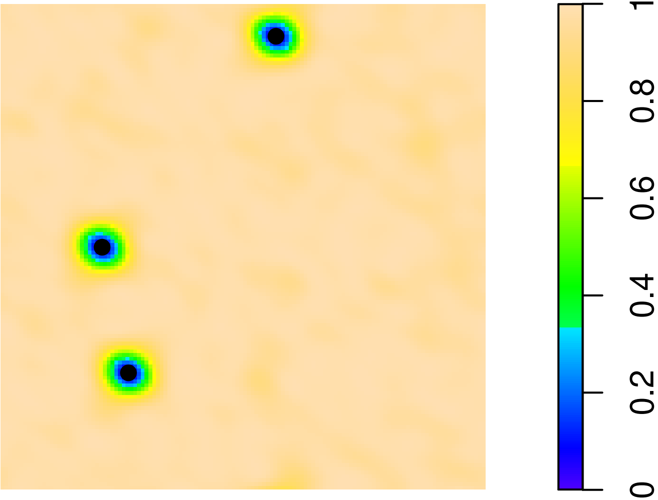
Step 3. The next point is sampled w.r.t. the following density:



Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

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



Dept. of Mathematical Sciences
Aalborg University
Denmark

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

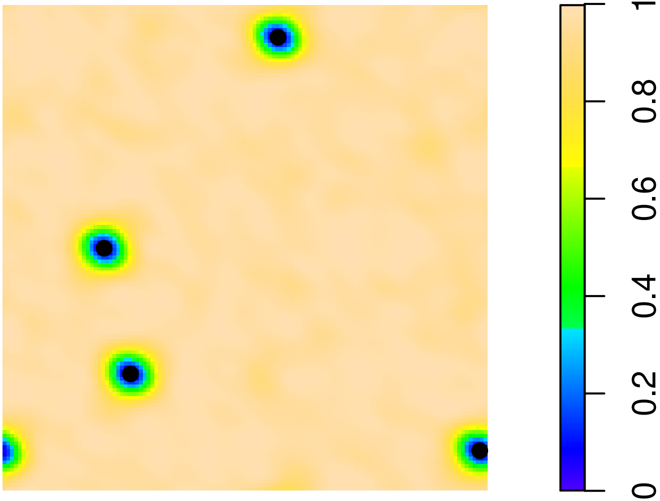
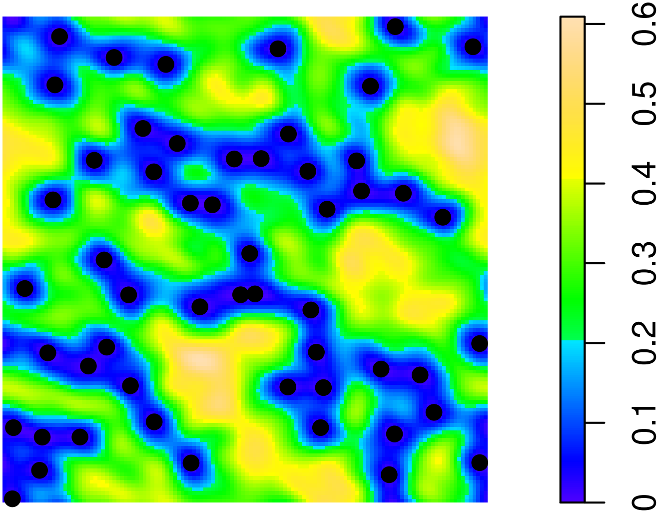


Illustration of simulation algorithm

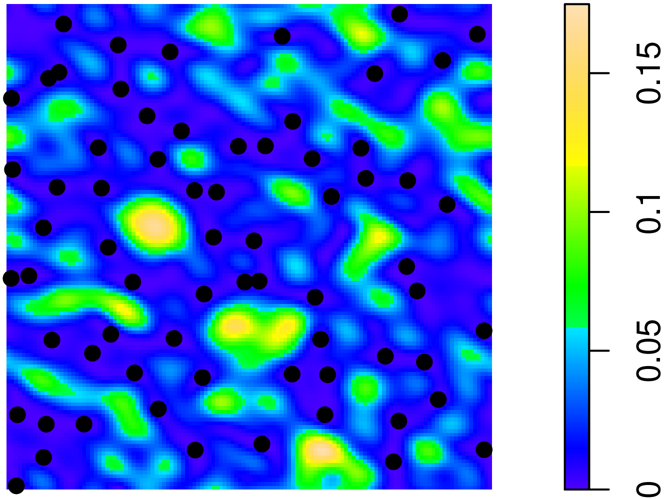
...somewhere in the middle...



Dept. of Mathematical Sciences
Aalborg University
Denmark

Illustration of simulation algorithm

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



Dept. of Mathematical Sciences
Aalborg University
Denmark

Simulation in R

Simply use the generic function `simulate`:

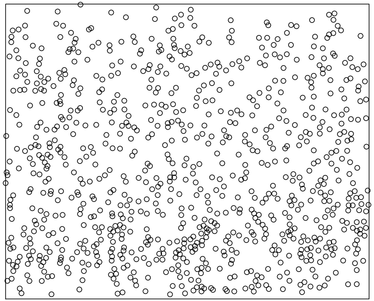
- ```
  model <- dppGauss(lambda=100, alpha=0.05, d=2)
  X <- simulate(model)
  ```
- Change the window (default is the unit square):
  ```
  W <- owin(poly=list(x=c(-1,0,1),y=c(0,1,0)))
  X <- simulate(model, W=W)
  ```
- Several realizations:
  ```
  X <- simulate(model, nsim=4)
  ```

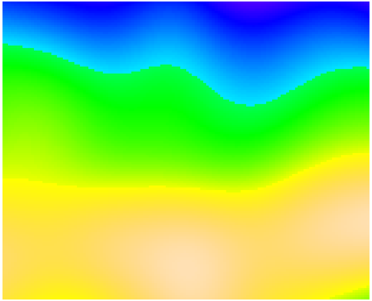
Ege Rubak

Dept. of Mathematical Sciences
Aalborg University
Denmark
  ```

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.

Ege Rubak

Dept. of Mathematical Sciences
Aalborg University
Denmark

Modelling inhomogeneity

Assume *correlation stationarity* (second-order intensity-reweighted stationarity) i.e. the correlation function is translation invariant:

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

such that $R(x,x) = 1$.

Modelling inhomogeneity

Assume *correlation stationarity* (second-order intensity-reweighted stationarity) i.e. the correlation function is translation invariant:

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

such that $R(x,x) = 1$.

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

Ege Rubak

Dept. of Mathematical Sciences
Aalborg University
Denmark

Modelling inhomogeneity

Assume *correlation stationarity* (second-order intensity-reweighted stationarity) i.e. the correlation function is translation invariant:

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

such that $R(x,x) = 1$.

- 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- og K-function.

Dept. of Mathematical Sciences
Aalborg University
Denmark

Assume *correlation stationarity* (second-order intensity-reweighted stationarity) i.e. the correlation function is translation invariant:

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

such that $R(x,x) = 1$.

- 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- og K-function.
- Fit a parametric model for R_0 via minimum contrast.

Dept. of Mathematical Sciences
Aalborg University
Denmark

Modelling inhomogeneity

Assume *correlation stationarity* (second-order intensity-reweighted stationarity) i.e. the correlation function is translation invariant:

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

such that $R(x, x) = 1$.

- 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- og K-function.
- Fit a parametric model for R_0 via minimum contrast.
- The resulting DPP has

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

Dept. of Mathematical Sciences
Aalborg University
Denmark