Dimension reduction in extended Quermass-interaction process

Kateřina Helisová

Czech Technical University in Prague

helisova@math.feld.cvut.cz

joint work with Jakub Staněk (Charles University in Prague)

20th January 2014

Outline

Outline

1. Point processes
2. Quermass-interaction process and its extension
3. Maximum likelihood method using MCMC
4. Dimension reduction

Point processes

Definition Consider N the system of locally finite subsets of $\mathbb{R}^d$ with the σ-algebra $\mathcal{N} = \sigma(\{\mathbf{x} \in N : \sharp(\mathbf{x} \cap A) = m\} : A \in \mathcal{B}, m \in \mathbf{N}_0)$. *A point process* X defined on $\mathbb{R}^d$ is a measurable mapping from some probability space $(\Omega, \mathcal{F}, P)$ to $(N, \mathcal{N})$.

Definition A locally finite, diffusion measure μ on $\mathcal{B}$ satisfying $\mu(A) = \mathrm{E}X(A)$ for all $A \in \mathcal{B}$ is called *the intensity measure*.

Definition If there exists a function $\rho(x)$ for $x \in \mathbb{R}^d$ such that $\mu(A) = \int_A \rho(x)dx$, then $\rho(x)$ is called *the intensity function*.

Definition If $\rho(x) = \rho$ is constant then the constant ρ is called *intensity*.

Back

Close

Poisson point process

Definition *The Poisson process* Y is the process which satisfies:

- for any finite collection $\{A_n\}$ of disjoint sets in $\mathbb{R}^d$, the numbers of points in these sets, $Y(A_n)$, are independent random variables,
- for each $A \subset \mathbb{R}^d$ such that $\mu(A) < \infty$, $Y(A)$ has Poisson distribution with parameter $\mu(A)$, i.e. $P[Y(A) = k] = \frac{\mu(A)^k}{k!} e^{-k}$, where μ is the intensity measure.

Point process given by the density with respect to Poisson process

Let Y be the Poisson process with an intensity measure μ.

For $F \in \mathcal{N}$, denote $\Pi(F) = P(Y \in F)$.

Definition A point process X is given by density f with respect to the Poisson process Y if

$$P(X \in F) = \int_F f(\mathbf{x})\Pi(d\mathbf{x}).$$

Outline

1. Point processes
2. Quermass-interaction process and its extension
3. Maximum likelihood method using MCMC
4. Dimension reduction

Notation

- $x = b(u, r)$ ... a disc with centre in $u \in \mathbb{R}^2$ and radius $r \in (0, \infty)$
- $\mathbf{x} = \{x_1, \ldots, x_n\}$ ... finite configuration of n discs
- $U_{\mathbf{x}}$ ... the union of discs from the configuration $\mathbf{x}$
- $\mathbf{Y}$ ... random disc Boolean model (i.e. union of discs without any interactions) with an intensity function of discs centers $\rho(u)$ and probability distribution of the discs radii Q
- $\mathbf{X}$ ... random disc process which is absolutely continuous with respect to the process $\mathbf{Y}$

Back
Close

Assumptions

- The intensity function is $\rho(u) = \rho > 0$ on a bounded set S and $\rho(u) = 0$ otherwise, i.e. the centers of the reference Boolean model form stationary Poisson process on S.

- For any finite configuration of discs $\mathbf{x} = \{x_1, \ldots, x_n\}$, the probability measure of $\mathbf{X}$ with respect to the probability measure of $\mathbf{Y}$ is given by density

$$f_\theta(\mathbf{x}) = \frac{\exp\{\theta \cdot T(U_\mathbf{x})\}}{c_\theta},$$

  where

  - c_θ is the normalizing constant,
  - θ is m-dimensional vector of parameters,
  - $T(U_\mathbf{x})$ is a m-dimensional vector of geometrical characteristics of the union $U_\mathbf{x}$ of the discs from the configuration $\mathbf{x}$.

Quermass-interaction process

The density is of the form

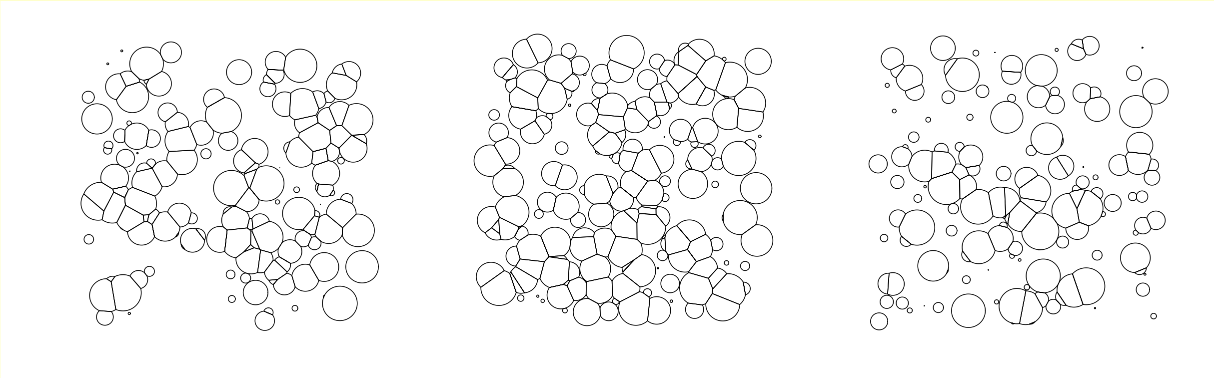
$$f_\theta(\mathbf{x}) = \frac{1}{c_\theta} \exp\{\theta_1 A(U_\mathbf{x}) + \theta_2 L(U_\mathbf{x}) + \theta_3 \chi(U_\mathbf{x})\},$$

where

- $A = A(U_\mathbf{x})$ is the area,
- $L = L(U_\mathbf{x})$ is the perimeter,
- $\chi = \chi(U_\mathbf{x})$ is the Euler-Poincaré characteristic (the number of connected components minus the number of holes, i.e. $\chi(U_\mathbf{x}) = N_{\mathrm{cc}}(U_\mathbf{x}) - N_{\mathrm{h}}(U_\mathbf{x})$)

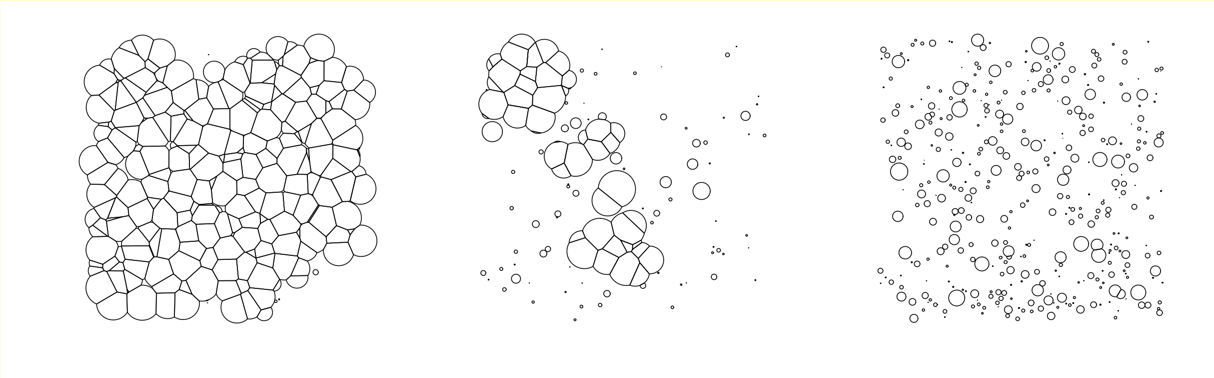
of the union $U_\mathbf{x}$.

Interpretation of the parameters



A realization of the reference random disc Boolean model on a rectangular region $S = [0, 30] \times [0, 30]$ with Q the uniform distribution on the interval $(0, 2)$ and $\rho = 0.2$ (left), and A-interaction model with parameters $\theta_1 = 0.1$ (middle), resp. $\theta_1 = -0.1$ (right).

Interpretation of the parameters



Quermass-interaction process with parameters $(\theta_1, \theta_2, \theta_3) = (0.6, -1, 1)$ (left), $(0.6, -1, 2)$ (middle) and $(0.6, -1, 5)$ (right).

Extended Quermass-interaction process

- Møller, Helisová (2008):
  - In the density

$$f_\theta(\mathbf{x}) = \frac{\exp\{\theta \cdot T(U_\mathbf{x})\}}{c_\theta},$$

    we have $T = (A, L, \chi, N_h, N_{bv}, N_{id})$, where
    $N_{bv} = N_{bv}(U_\mathbf{x})$ is the number of boundary vertices,
    $N_{id} = N_{id}(U_\mathbf{x})$ is the number of isolated discs
    of the union $U_\mathbf{x}$.
  - Theory and simulations studied.
- Møller, Helisová (2010):
  - $T = (A, L, N_{cc}, N_h)$.
  - Statistical analysis.

Outline

1. Point processes
2. Quermass-interaction process and its extension
3. Maximum likelihood method using MCMC
4. Dimension reduction

Maximum likelihood method using MCMC simulations (MCMC MLE)

- Denote $f_\theta(\mathbf{x}) = h_\theta(\mathbf{x})/c_\theta$ (i.e. $h_\theta(\mathbf{x}) = \exp\{\theta \cdot T(U_\mathbf{x})\}$ is the unnormalized density).

- For an observation $\mathbf{x}$, the log likelihood function is given by

$$l(\theta) = \log h_\theta(\mathbf{x}) - \log c_\theta = \theta \cdot T(U_\mathbf{x}) - \log c_\theta.$$

Problem 1: c_θ has no explicit expression.

Maximum likelihood method using MCMC simulations (MCMC MLE)

- Denote $f_\theta(\mathbf{x}) = h_\theta(\mathbf{x})/c_\theta$ (i.e. $h_\theta(\mathbf{x}) = \exp\{\theta \cdot T(U_\mathbf{x})\}$ is the unnormalized density).

- For an observation $\mathbf{x}$, the log likelihood function is given by

$$l(\theta) = \log h_\theta(\mathbf{x}) - \log c_\theta = \theta \cdot T(U_\mathbf{x}) - \log c_\theta.$$

Problem 1: c_θ has no explicit expression.

Solution of problem 1 (Møller, Waagepetersen (2004) applied by Møller, Helisová (2010)): We maximize the likelihood ratio f_θ/f_{θ_0} for a fixed vector θ_0 instead.

MCMC MLE - problem 1

Solution of problem 1: For the fixed θ_0, the log likelihood ratio

$$l(\theta) - l(\theta_0) = \log(h_\theta(\mathbf{x})/h_{\theta_0}(\mathbf{x})) - \log(c_\theta/c_{\theta_0})$$

can be approximated by

$$l(\theta) - l(\theta_0) \approx \log(h_\theta(\mathbf{x})/h_{\theta_0}(\mathbf{x})) - \log \frac{1}{M} \sum_{i=1}^{M} h_\theta(\mathbf{z}_i)/h_{\theta_0}(\mathbf{z}_i), \tag{1}$$

where $\mathbf{z}_i$, $i = 1, \ldots, M$, are realizations from f_{θ_0} obtained by MCMC simulations.

MCMC MLE - problem 2

Problem 2: MCMC MLE is time-consuming, because

- quite a large number M of realizations is needed,
- the approximation is possible only for θ_0 close to $\theta \Rightarrow$ bridge sampling ($\theta_0 \to \hat{\theta}^{(1)} \to \hat{\theta}^{(2)} \to \cdots \to \hat{\theta}$).

MCMC MLE - problem 2

Problem 2: MCMC MLE is time-consuming, because

- quite a large number M of realizations is needed,
- the approximation is possible only for θ_0 close to $\theta \Rightarrow$ bridge sampling $(\theta_0 \rightarrow \hat{\theta}^{(1)} \rightarrow \hat{\theta}^{(2)} \rightarrow \cdots \rightarrow \hat{\theta})$.

Solution of problem 2: Dimension reduction $\Rightarrow$ estimating θ by maximum likelihood method converts to looking for the maximum in lower-dimensional space.

MCMC MLE - problem 3

Problem 3: Let $U_{\mathbf{x}}$ be the observed set, $\mathbf{z}_i$, $i = 1, \ldots, M$, be the realizations from (1) and $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_m)$ be the corresponding MCMC maximum likelihood estimate of the parameter θ. Denote T_j the j-th item of the vector T. Then for all $j = 1, \ldots, m$, the following holds:

(i) If $T_j(U_{\mathbf{x}}) \leq T_j(U_{\mathbf{z}_i})$ for all $i = 1, \ldots, M$ and $T_j(U_{\mathbf{x}}) < T_j(U_{\mathbf{z}_i})$ for at least one i, then $\hat{\theta}_j = -\infty$.

(ii) If $T_j(U_{\mathbf{x}}) \geq T_j(U_{\mathbf{z}_i})$ for all $i = 1, \ldots, M$ and $T_j(U_{\mathbf{x}}) > T_j(U_{\mathbf{z}_i})$ for at least one i, then $\hat{\theta}_j = \infty$.

MCMC MLE - problem 3

Problem 3: Let $U_{\mathbf{x}}$ be the observed set, $\mathbf{z}_i$, $i = 1, \ldots, M$, be the realizations from (1) and $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_m)$ be the corresponding MCMC maximum likelihood estimate of the parameter θ. Denote T_j the j-th item of the vector T. Then for all $j = 1, \ldots, m$, the following holds:

(i) If $T_j(U_{\mathbf{x}}) \leq T_j(U_{\mathbf{z}_i})$ for all $i = 1, \ldots, M$ and $T_j(U_{\mathbf{x}}) < T_j(U_{\mathbf{z}_i})$ for at least one i, then $\hat{\theta}_j = -\infty$.

(ii) If $T_j(U_{\mathbf{x}}) \geq T_j(U_{\mathbf{z}_i})$ for all $i = 1, \ldots, M$ and $T_j(U_{\mathbf{x}}) > T_j(U_{\mathbf{z}_i})$ for at least one i, then $\hat{\theta}_j = \infty$.

Solution of problem 3: Dimension reduction $\Rightarrow$ transformation of used geometrical characteristics.

Outline

1. Point processes
2. Quermass-interaction process and its extension
3. Maximum likelihood method using MCMC
4. Dimension reduction

Principal components method

- Denote $\mathbf{V} = (\sigma^2_{i,j})^m_{i,j=1}$ the variance matrix of $T(U_\mathbf{x})$.
- Suppose that $\mathbf{V}$ has $r \geq 0$ positive, mutually different eigenvalues $\lambda_1 > \lambda_2 > \ldots > \lambda_r \to$ corresponding eigenvectors $\mathbf{v}_1$,...,$\mathbf{v}_r$.
- Looking for a vector $\mathbf{u}$ such that $\mathbf{u}^T\mathbf{u} = 1$ and $\mathbf{u}T(U_\mathbf{x})$ has the largest possible variance $\to \mathbf{u} = \mathbf{v}_1$ & $var(\mathbf{v}_1 T(U_\mathbf{x})) = \lambda_1$.
- Denote $C_1(U_\mathbf{x}) = \mathbf{v}_1 T(U_\mathbf{x})$.
- Looking for a vector $\mathbf{u}$ such that $\mathbf{u}^T\mathbf{u} = 1$, $\mathbf{u}T(U_\mathbf{x})$ has the largest possible variance and $cov(\mathbf{u}T(U_\mathbf{x}), C_1(U_\mathbf{x})) = 0 \to \mathbf{u} = \mathbf{v}_2$.
- Denote $C_2(U_\mathbf{x}) = \mathbf{v}_2 T(U_\mathbf{x}) \to var C_2(U_\mathbf{x}) = \lambda_2$.
- $C_1(U_\mathbf{x}), \ldots, C_r(U_\mathbf{x}) \to$ principal components of the vector $T(U_\mathbf{x})$.
- $\mathbf{v}_1(U_\mathbf{x}), \ldots, \mathbf{v}_r(U_\mathbf{x}) \to$ principal directions.

Principal components method

- Usually in practice $r = m$.
- Denoting

$$\sigma^2 = \sum_{i=1}^{m} \sigma_{ii}^2,$$

it can be proved that $varC_1 + \ldots + varC_r = \lambda_1 + \ldots + \lambda_r = \sigma^2$.

$$\Downarrow$$

For $p < r$, $C_1(U_{\mathbf{x}}), \ldots, C_p(U_{\mathbf{x}})$ such that

$$\frac{\lambda_1 + \ldots + \lambda_p}{\sigma^2} \doteq 1$$

cover the variability of the data enough and can explain the behaviour of the vector $T(U_{\mathbf{x}})$ satisfactorily.

Principal components method

- Our aim: rewrite the density

$$f_\theta(\mathbf{x}) = \frac{\exp\{\theta \cdot T(U_\mathbf{x})\}}{c_\theta} = \frac{\exp\{\theta_1 T_1(U_\mathbf{x}) + \ldots + \theta_m T_m(U_\mathbf{x})\}}{c_\theta}$$

  to the form

$$f_\varphi(\mathbf{x}) = \frac{\exp\{\varphi \cdot C(U_\mathbf{x})\}}{c_\varphi} = \frac{\exp\{\varphi_1 C_1(U_\mathbf{x}) + \ldots + \varphi_p C_p(U_\mathbf{x})\}}{c_\varphi},$$

  where

  – φ has lower dimension than $\theta \Rightarrow$ its estimation is faster,

  – items of $C(U_\mathbf{x})$ can be both positive and negative $\rightarrow$ no undervaluation or overvaluation of parameter estimates.

Principal components method

- Our aim: rewrite the density

$$f_\theta(\mathbf{x}) = \frac{\exp\{\theta \cdot T(U_\mathbf{x})\}}{c_\theta} = \frac{\exp\{\theta_1 T_1(U_\mathbf{x}) + \ldots + \theta_m T_m(U_\mathbf{x})\}}{c_\theta}$$

to the form

$$f_\varphi(\mathbf{x}) = \frac{\exp\{\varphi \cdot C(U_\mathbf{x})\}}{c_\varphi} = \frac{\exp\{\varphi_1 C_1(U_\mathbf{x}) + \ldots + \varphi_p C_p(U_\mathbf{x})\}}{c_\varphi},$$

where

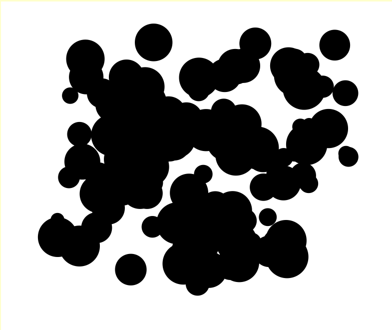
  - φ has lower dimension than $\theta \Rightarrow$ its estimation is faster,
  - items of $C(U_\mathbf{x})$ can be both positive and negative $\rightarrow$ no undervaluation or overvaluation of parameter estimates.

- Different ways how to determine p described (Rencher 2002), e.g. to take such p that the cumulative variance $(\lambda_1 + \ldots + \lambda_p)$ is greater than 80% of total variance $(\lambda_1 + \ldots + \lambda_r)$.

Numerical results - simulated data

N=100 realizations of the $(A, L, N_{cc}, N_h, N_{id})$-interaction process with

- centers in 10×10 square window,
- parameters $\theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) = (1.5, -1, 1, -0.25, -0.5)$,
- reference process with
  - the intensity of the disc centers $\rho = 1$,
  - radii of discs uniformly distributed in the interval $[0.2, 0.7]$.



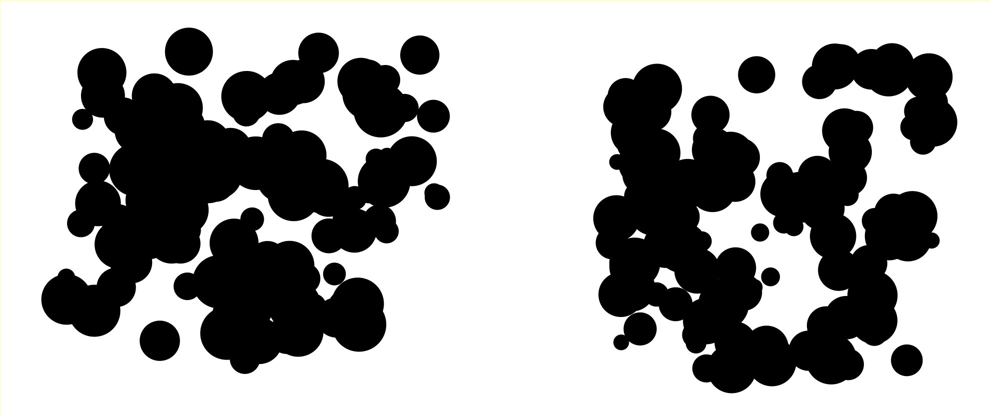
Numerical results - simulated data

N	eigenvalues	corresponding eigenvectors	cumul.var
100	**117.40**	**(-0.41,0.82,0.28,-0.25,0.11)**	49%
	104.15	**(-0.77,-0.55,0.22,-0.23,0.05)**	93%
	8.35	(0.06, -0.02,0.64,0.63,0.44)	97%
	5.71	(0.43,-0.15,0.30,-0.69,0.47)	99%
	1.57	(0.20,-0.05,0.61,-0.11,-0.75)	100%

Numerical results - simulated data

N	eigenvalues	corresponding eigenvectors	cumul.var
100	**117.40**	**(-0.41,0.82,0.28,-0.25,0.11)**	49%
	104.15	**(-0.77,-0.55,0.22,-0.23,0.05)**	93%
	8.35	(0.06, -0.02,0.64,0.63,0.44)	97%
	5.71	(0.43,-0.15,0.30,-0.69,0.47)	99%
	1.57	(0.20,-0.05,0.61,-0.11,-0.75)	100%
10	**98.53**	**(0.61,-0.47,-0.52,0.29,-0.21)**	63%
	46.70	**(0.61,0.78,0.01,0.09,0.11)**	93%
	7.16	(0.06,0.09,-0.43,-0.88,-0.19)	98%
	2.58	(-0.49,0.39,-0.60,0.37,-0.34)	99%
	0.78	(-0.10,-0.04,-0.43,0.02,0.89)	100%

Model checking



Comparing a realization of the original data (left) with a realization of fitted model for $N = 10$ input sets (right).

Model checking - CDF

- For a random set $\mathbf{Z}$ and a compact convex set $B \subset \mathbb{R}^2$, define

$$D = inf\{r \geq 0 : \mathbf{Z} \cap rB \neq \emptyset\}.$$

If $P(D > 0) > 0$ and B is the unit disc $b(0,1)$, then the spherical contact distribution function of the random set $\mathbf{Z}$ is defined as

$$H_B(r) = P(D \leq r | D > 0).$$

- Estimator for stationary $\mathbf{Z}$:

$$\hat{H}_B(r) = \frac{\sum_{u \in G} \mathbb{I}_{[u \notin \mathbf{Z},\, u+rB \subset W,\, (u+rB) \cap \mathbf{Z} \neq \emptyset]}}{\sum_{u \in G} \mathbb{I}_{[u \notin \mathbf{Z},\, u+rB \subset W]}}$$

where G is a lattice in observation window W.

Model checking - covariance

- The covariance function of a motion invariant (i.e. stationary and isotropic) random set $\mathbf{Z}$ is defined as

$$C(r) = P(u \in \mathbf{Z}, v \in \mathbf{Z}),$$

  where $\|u - v\| = r$.

- Estimator for motion invariant $\mathbf{Z}$:

$$\hat{C}(r) = \frac{\sum_{u,v \in G} \mathbb{I}_{[\|u-v\|=r,\, \{u,v\} \subset \mathbf{Z}]}}{\sum_{u,v \in G} \mathbb{I}_{[\|u-v\|=r]}}$$

  provided the denominator is non-zero.

Model checking - shape characteristics

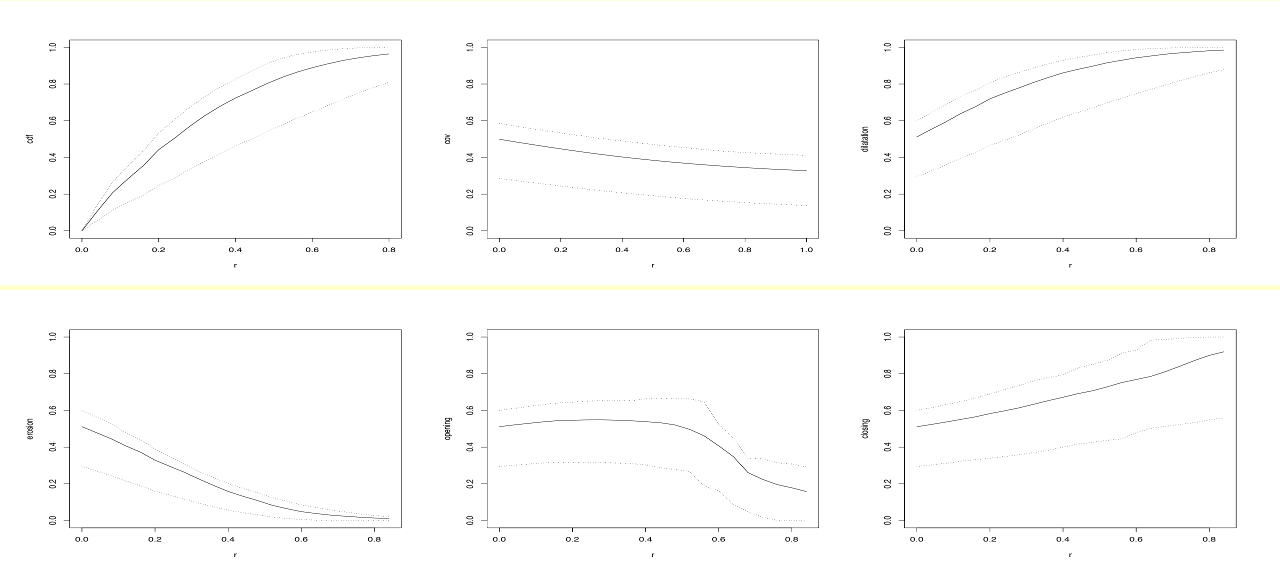
- Denote
  $|\mathbf{Z}| = A(\mathbf{Z})$,
  $\mathbf{Z}_{\ominus r} = \{u \in \mathbb{R}^2 : b(u,r) \subseteq \mathbf{Z}\}$ for $r > 0$,
  $\mathbf{Z}_{\oplus r} = \cup_{u \in \mathbf{Z}} b(u,r)$ for $r > 0$.

- Dilation d, erosion e, opening o and closing c of $\mathbf{Z}$ by the disc $b(0,r)$ are defined by

$$d(r) = \frac{|\mathbf{Z}_{\oplus r} \cap W_{\ominus r}|}{|W_{\ominus r}|}, \quad e(r) = \frac{|\mathbf{Z}_{\ominus r}|}{|W_{\ominus r}|},$$

$$o(r) = \frac{|(\mathbf{Z}_{\ominus r})_{\oplus r} \cap W_{\ominus 2r}|}{|W_{\ominus 2r}|}, \quad c(r) = \frac{|(\mathbf{Z}_{\oplus r})_{\ominus r} \cap W_{\ominus 2r}|}{|W_{\ominus 2r}|}.$$

Model checking



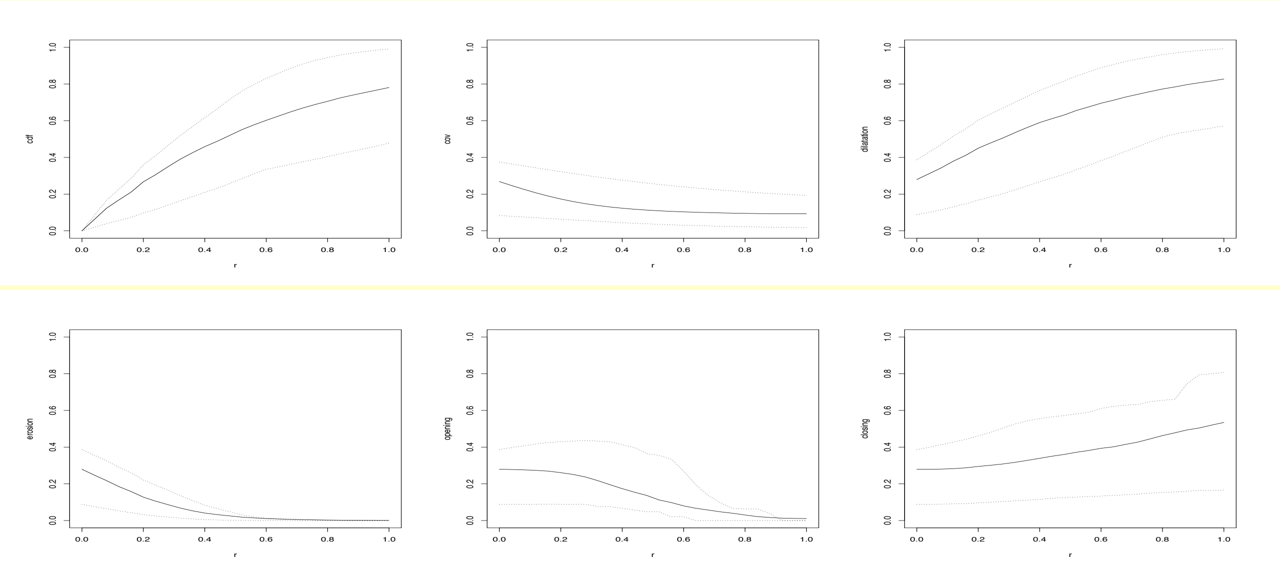
Contact distribution function, covariance function, dilatation, erosion, opening and closing averaged from 10 input realizations (full lines) and 95%-envelopes build from 39 simulations of the fitted model.

Numerical results - real data

Left: Data from Mrkvička, Mattfeldt (2011) - cells of mammary cancer.

Right: Realization of fitted model obtained by principal component method using 10 sets on the input - two principal directions appeared significant.

Numerical results - real data



Contact distribution function, covariance function, dilatation, erosion, opening and closing averaged from 10 input realizations (full lines) and 95%-envelopes build from 39 simulations of the fitted model.

References

- Møller J., Helisová K. (2008): Power diagrams and interaction process for unions of discs. Advances in Applied Probability 40(2), 321–347.
- Møller J., Helisová K. (2010): Likelihood inference for unions of interacting discs. Scandinavian Journal of Statistics 37(3), 365–381.
- Møller J., Waagepetersen R.P. (2004): Statistical inference and simulations for spatial point processes. Chapman and Hall/CRC, Boca Raton.
- Mrkvička T., Mattfeldt T. (2011): Testing histological images of mammary tissues on compatibility with the Boolean model of random sets. Image Analysis & Stereology 30(1), 11–18.
- Rencher A.C. (2002): Methods of Multivariate Analysis, 2nd edn. Wiley & Sons, New York.
- Staňková Helisová K., Staněk J. (2012): Dimension reduction in extended Quermass-interaction process. Methodology and Computing in Applied Probability. DOI 10.1007/s11009-013-9343-x.

Thank you for your attention!