Sequential Point Process Model and Bayesian Inference for Spatial Point Patterns with Linear Structures

Jakob G. Rasmussen

Joint work with Jesper Møller

Department of Mathematical Sciences
Aalborg University
Denmark

Outline

Data

- Dataset 1: Barrows
- Dataset 2: Mountain tops

Model

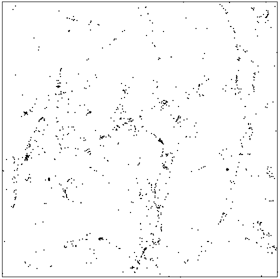
- Model construction
- Simulation algorithm

Inference

- Bayesian model: likelihood and priors
- MCMC based parameter estimation

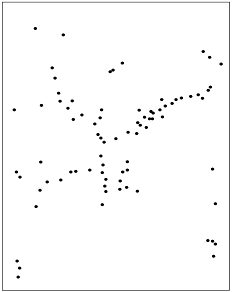
Dataset 1: Barrows

- A barrow is a bronze age burial site resembling a small hill.
- These are important sources of information for archaologists.
- They are often placed roughly in linear structures.



Dataset 2: Mountain tops

- Mountains ridges means that "local" tops are often forming linear structures.



Linear structures

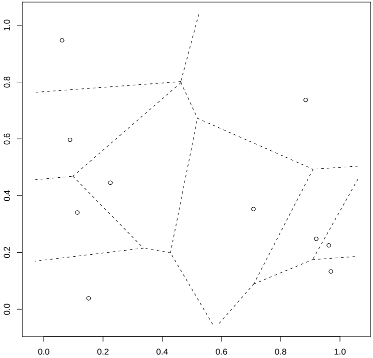
- In this talk we will consider a model capable of generating linear formations.
- Roughly speaking, this model generates linear structures by moving points closer to other points.
- Interpretation of the model:
  - Barrows: Here the model is interpreted as dead people are buried close to previously buried people.
  - Mountains: No reasonable interpretation - the model should not be thought of as representing actual mechanics.

Outline

Model construction

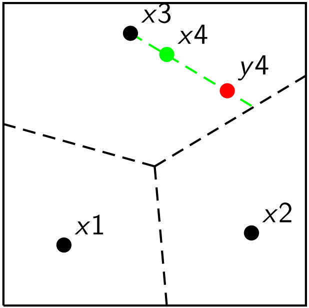
- Point process x defined on window W.
- $x = x_c \cup x_b$ with n points.
- Number of points in x_c, k, is binom(n, q).
- Background process:
  - x_b consists of i.i.d. uniformly distributed points on W
- Cluster process:
  - Sequential construction.
  - A point is initially uniformly distributed independently of everything else.
  - With probability p this point is moved closer to the closest previous point; otherwise it keeps its original position.

Voronoi tesselations



- Voronoi tesselation: an area is associated with the closest point in the point process

Moving points



- Density for new position:

$$h(x_i|\{x_1,\ldots,x_{i-1}\};\sigma^2) \propto \exp(-r_i^2/(2\sigma^2)), \qquad 0 < r_i < l_i$$

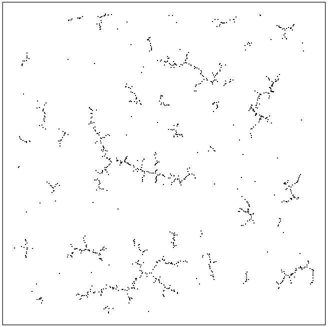
- Other distributions have also been tried.

Simulation algorithm

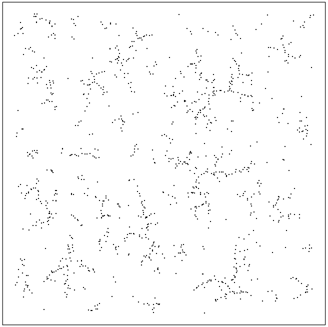
Fix number of points, and for each point do the following:

1. Find type of point i (cluster with prob. q, background otherwise)
2. If background point:
   2.1 Find coordinates - uniformly distributed on W
3. If cluster point:
   3.1 Find initial coordinates - uniformly distributed on W
   3.2 Move with probability p, otherwise keep position
   3.3 If move, find closest cluster point and move new point closer to this using Exp-distribution

Two simulations



$q = 0.95, p = 0.95, \sigma = 70$



$q = 0.90, p = 0.90, \sigma = 130$

Remember:

- q is the probability that a point is a cluster point (i.e. belongs to a linear structure)
- p is the probability is that a cluster point is moved (i.e. continues an existing linear structure)
- σ governs how close points in lines are located to each other

Outline

Likelihood and priors

- Let z be the observed point pattern x including type (cluster/background) and order of points.
- Likelihood:

$$L(q, p, \sigma^2 | z) = \binom{n}{k} q^k \left(\frac{1-q}{|W|} \right)^{n-k} \prod_{i=1}^{k} f(x_i | x_1, \ldots, x_{i-1}; p, \sigma^2)$$

  where

$$f(\cdot | x_1, \ldots, x_{i-1}; p, \sigma^2) = p \times h(\cdot | \{x_1, \ldots, x_{i-1}\}; \sigma^2) + (1-p) \times \frac{1}{|W|}$$

- Priors:
  - Independent priors for p, q, σ.
  - p, q: Uniform on $[0, 1]$.
  - σ: Flat inverse gamma or (improper) uniform on $[0, \infty)$.

Estimation of parameters

- Ideally we would find mean/maximum of the posterior

$$p(q, p, \sigma^2 | x) \propto g_1(p) g_2(q) g_3(\sigma^2) L(q, p, \sigma^2 | x)$$

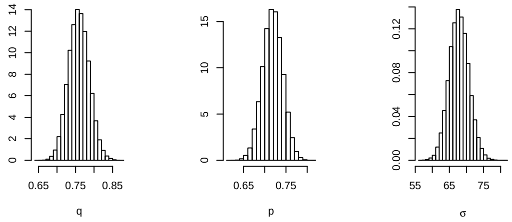
  but we only have closed form expression for $L(q, p, \sigma^2 | z)$, not $L(q, p, \sigma^2 | x)$.
- So we have a missing data problem:
  - The order of $x_c = \{x_1, \ldots, x_k\}$ is unknown.
  - Also it is unknown whether a point belongs to x_c or x_b.
- So we need to approximate the estimates of p, q, σ and the missing data by Markov chain Monte Carlo.

MCMC

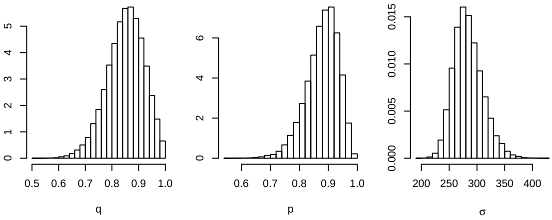
- We use Metropolis within Gibbs to obtain posterior.
- Updates:
  - A background point becomes a cluster point.
  - A cluster point becomes a background point.
  - Shifting the ordering of two succeeding cluster points.
  - Parameters p, q and σ^2: Metropolis update, normal proposal.
- Hastings ratios are easily obtained.
- Calculation times are not too bad.
- Mixing is fair.

Posterior distributions - parameters

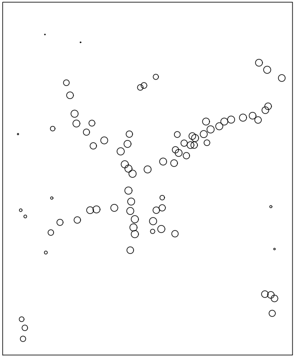
Barrows:

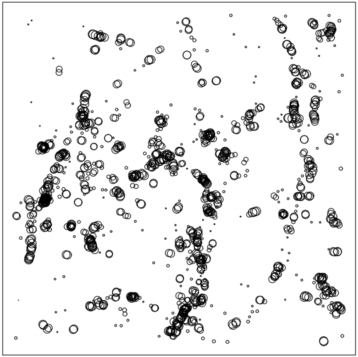


Mountains:



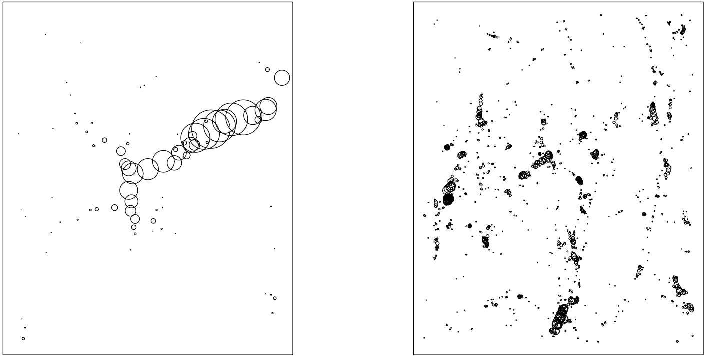
Posterior distributions - missing data





Circle radius indicates marginal posterior probability of a point being a cluster point.

Posterior distributions - missing data



Circle radius indicates the order in which the cluster points occur.

Concluding remarks

- Summing up: a new model with linear structures and MCMC-based Bayesian inference
- Model checking skipped in this talk
- Many extensions/modifications possible, e.g. inclusion of covariates, initial placements or moving mechanims

Thank you for your attention :-)