Point processes on directed linear networks

Jakob G. Rasmussen
Department of Mathematics
Aalborg University
Denmark

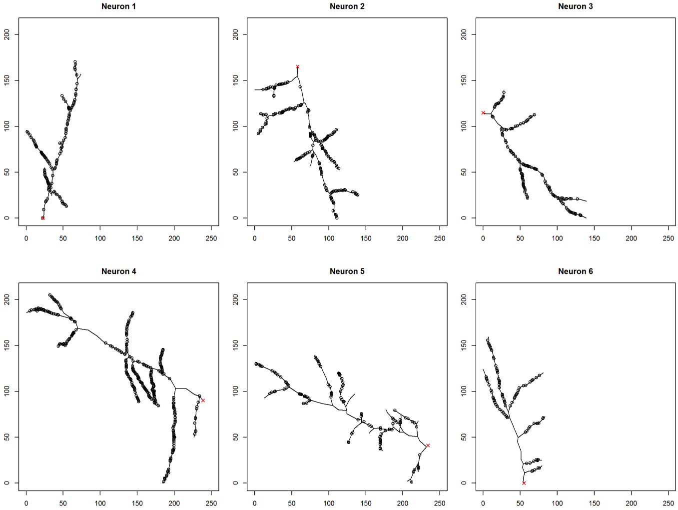
Contents

Why directed linear networks

- You have seen point defined on linear networks.
- Some linear networks have directions associated to each line segment, for example:
  - River networks, direction of the flow of water
  - Dendrite trees, direction away from root
- Potentially a point pattern observed on a directed linear network may have somehow evolved according to these directions.
- We can extend the definition of the conditional intensity function to directed linear networks, and get similar results to the ones obtained for temporal point processes, e.g. likelihood estimation, simulation and residual analysis.

Dendrite data

- 6 dendrite trees extending from neurons with spines
- Red cross is root

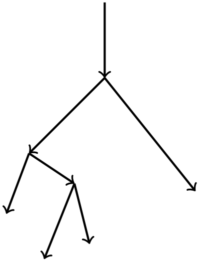
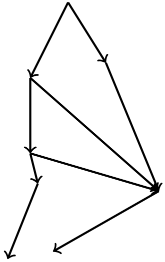
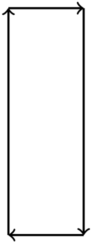


Exercise

1. As a warm up, try to come up with hypothetical data example(s) with
   - a linear network with directions, and
   - point pattern data observed on this network.
2. In your example, would you expect the directions to have and influence on the development of the points?

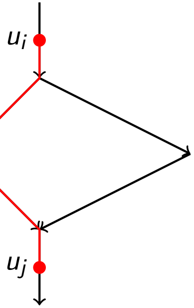
Directed linear networks

- Definition of DLN:
  - L is the union of line segments L_i, $i = 1 \ldots, n$
  - Each L_i has an associated direction
- DALN: Acyclic, meaning no cycle of L_i with same directions.
- Direction relation: $L_i \to L_j$ if there is a path following directions from L_i to L_j (similar meaning of $u_i \to u_j$, for $u_i, u_j \in L$). We say L_i is upstream of L_j, or L_j is downstream of L_i.

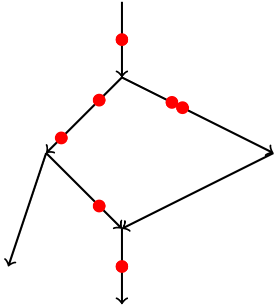
Shortest directed path distance

- Shortest directed path distance $d_L^{\rightarrow}$:
  - $d_L^{\rightarrow}(u_i, u_j)$ is the length of the shortest path following directions
  - Infinite, if no directed path exists
  - Quasimetric ($d_L^{\rightarrow}(u, u) = 0$, $d_L^{\rightarrow}(u, v) > 0$ for $u \neq v$, triangle inequality - but not symmetric)



Point patterns/processes on directed linear networks

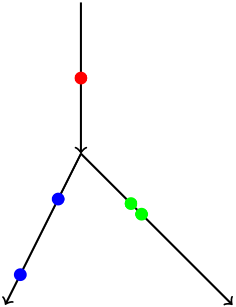
- Point pattern: A finite set of points $x_1, \ldots, x_n \in L$.
- Point process: A stochastic process with point patterns on L as realizations. Here we can use either the definition using counting measures or random locally finite sets.

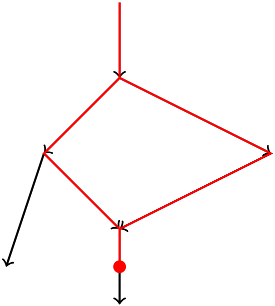


- Marked case: Here additional values (marks) are associated to each point, say x_i has mark m_i.

Point processes obeying directions

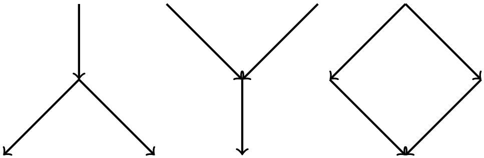
- The definition of the points on a directed linear network includes the cases not incorporating directions, but we will now focus on those that do.
- Imagine some random mechanism where points appear first at the top of the network, and then progressively further down the network.
- Whether or not a points appears at $u \in L$ may depend on all points upstream $L_{\to u} = \{v \in L : v \to u\}$, but not downstream.





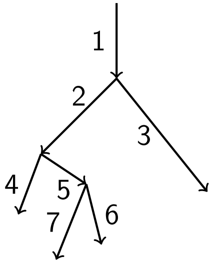
Exercise

1. In the left network, are the point processes on the two lower line segments independent? Are they conditionally independent given the point process on the upper line segment?
2. In the middle network, answer the same question for the two upper line segments (in the conditional case, condition on the lower line segment.)
3. In the right network, are any of point processes on the various line segments independent?



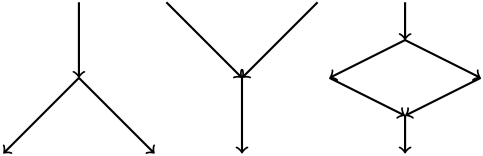
Conditional intensity functions

- We define a point process segment-wise using a conditional intensity function.
- Conditional intensity function: $\lambda^*(u)$ for $u \in L_i$ is the conditional intensity for a line segment associated to a time interval of the same length and direction, which may depend on all points in $L_{\to u}$.
- Note that for a DALN there exists a (non-unique) order $\omega \in \Omega$ of all $L_i \in L$, such that all $L_j \in L_{\to L_i}$ comes before L_i for every i. This means that we can assign λ^* to L one segment at a time.



Conditional intensity functions

- Marked case: Same definition of $\lambda^*(u, m)$, but with a mark m. Note that $\lambda^*(u, m) = \lambda^*(u) f^*(m|u)$, where $\lambda^*(u)$ is the ground intensity and $f^*(m|u)$ is the conditional mark distribution.
- New point processes on DALNs are specified by specifying a mathematical expression for λ^*.
  - In practice we need to specify it meaningfully for the whole network, even though it is technically defined for each line segment separately.
  - The tricky point is to specify a model such that it behaves meaningfully around junctions and line segments that are met on multiple upstream paths.

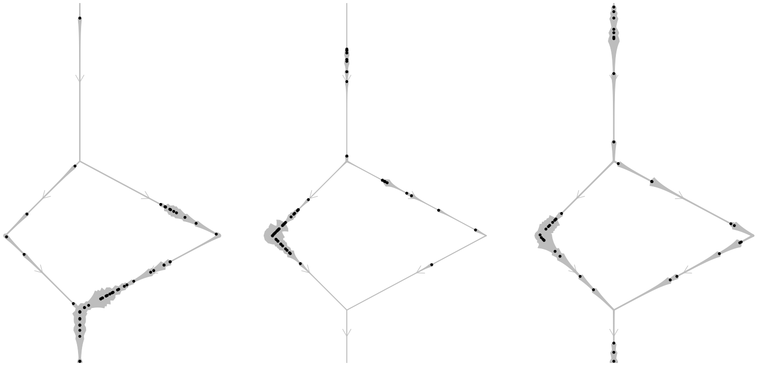


Example: Poisson process

- If $\lambda^*(u)$ does not actually depend on the upstream points, i.e. $\lambda^*(u) = \lambda(u)$ is a deterministic function, then the process is a Poisson process on L.
- In the homogeneous case, $\lambda^*(u) = \lambda$, the parameter λ is the mean number of points per unit length of line segments.
- The Poisson process on a directed linear network is equivalent to the Poisson process defined on the same network without directions.

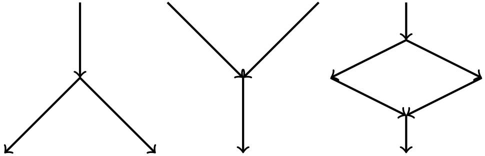
Example: Hawkes process

- CIF: $\lambda^*(u) = \mu + \alpha \sum_{x_i \in \boldsymbol{x} \colon x_i \to u} \gamma(d_L^{\to}(x_i, u))$
- Parameters: $\mu, \alpha \in [0, \infty)$, $\gamma(\cdot)$ is a density on $[0, \infty)$.
- This produces clustered point patterns.
- Three simulations with (μ, α, κ) equal to $(1, 0.8, 5)$, $(1, 0.8, 10)$, $(1, 0.9, 5)$, where $\gamma(t) = \kappa \exp(-\kappa t)$.



Behavior of the Hawkes process on L

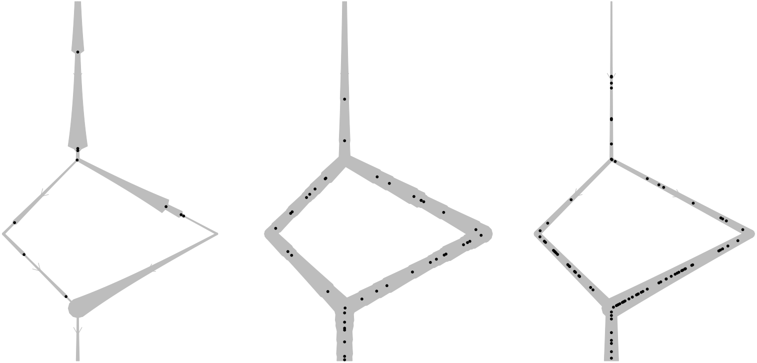
- The Hawkes process can be interpreted with the usual clustering and branching structure but
- For a splitting connections, a cluster will be copied to each outgoing line segment, resulting in more points - α still controls the number of offspring, but it is not the mean number of offspring anymore.
- Exercise: Consider the last two structures below. How does the clusters behave on these structures?



- The Hawkes process does have some weird behavior when we interpret the clusters in details, but if we just view it as a clustered point process, it is a fine model. Also to get specific behavior we can modify λ^* depending on L.

Example: Selfcorrecting process

- CIF: $\lambda^*(u) = \exp\{\mu d_L^{\rightarrow}(v_0, u) - \alpha|\boldsymbol{x} \cap \mathrm{sp}(v_0, u)|\}$
- Parameters: $\mu, \alpha \in [0, \infty)$
- $|\boldsymbol{x} \cap \mathrm{sp}(v_0, u)|$ is the number of points on the shortest path from unique root v_0 to x.
- This produces regular point patterns.
- Three simulations with $(\mu, \alpha) = (0.8, 1)$, $(0.4, 0.1)$, $(0.3, 0)$.



Example: Marked Hawkes process

- We can add marks to a Hawkes process, for example to get a multitype version, $m \in \{1, \ldots, K\}$.
- CIF:

$$\lambda^*(u, m) = \mu_m + \sum_{x_i \in \boldsymbol{x}:\, x_i \to u} \alpha_{m_i,m} \gamma_{m_i,m}(d_L^{\to}(x_i, u))$$

- This produces K clustered inter-dependent point patterns, which also has a tendency to cluster together.
- In the dendrite example, the spines on the dendrite actually have four different types, so the multitype Hawkes process might be relevant for a full analysis of this.

Maximum likelihood estimation

- Likelihood function for observed point pattern $\boldsymbol{x}$:

$$\mathcal{L}(\boldsymbol{\theta};\boldsymbol{x}) = \left(\prod_{x\in\boldsymbol{x}} \lambda^*(x;\boldsymbol{\theta})\right) \exp\left(-\int_L \lambda^*(u;\boldsymbol{\theta})\mathrm{d}\lambda_1(u)\right)$$

- Likelihood function for observed marked point pattern $\boldsymbol{y}$:

$$\mathcal{L}(\boldsymbol{\theta};\boldsymbol{y}) = \left(\prod_{(x,m)\in\boldsymbol{y}} \lambda^*(x,m;\boldsymbol{\theta})\right) \exp\left(-\int_L \lambda^*(u;\boldsymbol{\theta})\mathrm{d}\lambda_1(u)\right)$$

  For the marked case $\lambda^*(u;\boldsymbol{\theta})$ is the ground intensity.
- Homogeneous Poisson process: Just as for the temporal case, we get

$$\mathcal{L}(\boldsymbol{\theta};\boldsymbol{x}) = \lambda^{n(\boldsymbol{\theta})} e^{-\lambda|L|} \quad \text{and} \quad \hat{\lambda} = \frac{n(\boldsymbol{\theta})}{|L|}$$

- For the Hawkes and selfcorrecting processes, we can easily obtain fairly nice expressions for the likelihoods, but they cannot be maximized analytically, so we use approximate maximization, e.g. Newton-Raphson.

Simulation

- The point pattern $\boldsymbol{x}$ is simulated on one line segment at a time following an order $\omega \in \Omega$.
- Each $\boldsymbol{x}_i \in L_i$ is simulated conditional on $x_{\to i}$ using a adapted version of the following:
  1. Ogata's modified thinning algorithm.
  2. Inverse method.

  Both methods work on each line segment in exactly the same way as for the temporal point process.
- In the marked case, we simulate each mark using $f^*(m|u)$, also in the same way as for the marked temporal point process.
- In the simulations in this lecture, the inverse method has been used, but Ogata's method would work just as well.

Model checking: Residual analysis

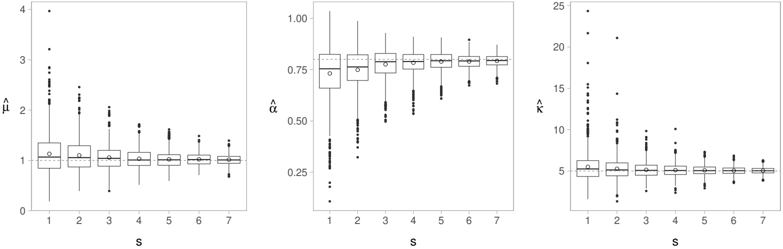
- Assume estimated model with integrated intensity $\hat{\Lambda}^*$ obtained from data.
- Residual process: Use $\hat{\Lambda}^*$ to transform the point patterns on each line segment following an order $\omega \in \Omega$.
- If model is good, the residual process should look like a unit-rate Poisson process - techniques for checking that a process is a homogeneous Poisson process can then be applied. This is more complicated than before, since this is a Poisson process on a linear network.
- Note that since length of line segments also change with transformation, we do not get a Poisson process on a transformed network; in fact, the obtained process may not be observed on a linear network that follows normal geometry.
- The marked case is difficult. For example, for the multitype process we get different networks for each type.

Data analysis: Simulated data

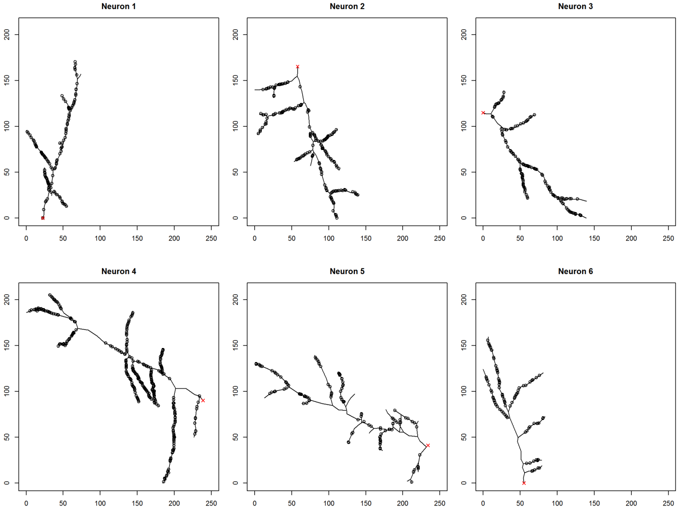
- Question: How well can we recover the parameters from simulated data using MLE?
- A Hawkes process has been simulated with parameters $(\mu, \alpha, \kappa) = (1, 0.8, 5)$ and γ is the density of an exponential distribution with parameter κ.
- Seven copies of the same network have been used, each increased by 50% since the last one.

Data analysis: Simulated data

- 1000 simulations has been made on each of variously sized networks, and MLE has been used to estimate the parameters.
- The following shows box-plots of the estimated parameters.
- The estimates seem fine, and improves with increasing network size.

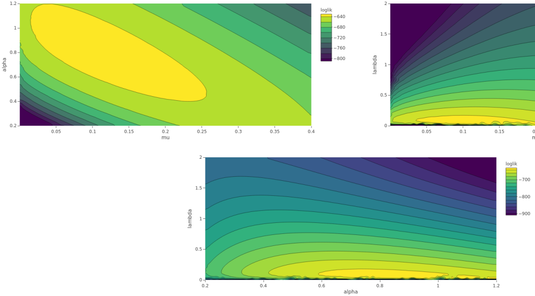


Data analysis: Dendrite data, estimation



- Data looks clustered - we model neuron 4.
- Model: Hawkes process with exponential γ.

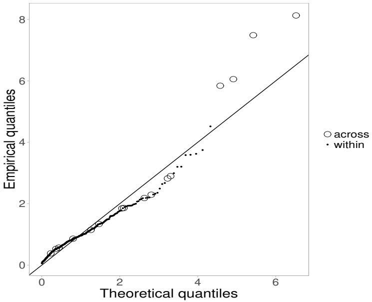
Data analysis: Dendrite data, estimation



- Plots of the likelihood as a function of (μ, α), (μ, κ) and (α, κ) with the remaining parameter fixed at the MLE.
- MLEs for neuron 4 given by $(\hat{\mu}, \hat{\alpha}, \hat{\kappa}) = (0.11, 0.84, 0.073)$

Data analysis, dendrite data, model checking

- Calculate residuals using $\hat{\Lambda}^*$ to obtain a new point pattern on a new network - should resemble a unit rate Poisson process.
- Calculate interevent distances.
- QQ-plot, residuals vs theoretical quantiles from Exp(1)-distribution.
- "Across" means interevent distances across junctions, "within" are the rest



Concluding remarks

- More details in:
  Rasmussen, J. G., Christensen, H. S. Point processes on directed linear networks. *Methohodology and Computing in Applied Probability*, volume 23, pages 647–667 (2021)
- The methodology and models for point processes on directed linear networks are rather new and a lot of questions are still unanswered.