Asymptotics for Euclidean minimal spanning trees on random points by David Aldous and J. Michael Steele

Bharath Roy Choudhury

ENS-PSL
INRIA

19 February 2020

Euclidean minimal spanning tree (MST)

Let $X = \{x_1, x_2, ..., x_n\}$ be points in $\mathbb{R}^d$ ($d \geq 2$). The minimal spanning tree (MST) t with vertex set X is a tree whose sum of the edge-lengths are minimum.

$$\sum_{e \in t} |e| = \min_G \sum_{e \in G} |e|$$

Goal of the authors in this paper:

Study expectation of functionals on the MST on Euclidean random points obtained by a Poisson point process.

Goal of the authors in this paper:

Study expectation of functionals on the MST on Euclidean random points obtained by a Poisson point process.

What are these functionals?

- Degree of a vertex.

Goal of the authors in this paper:

Study expectation of functionals on the MST on Euclidean random points obtained by a Poisson point process.

What are these functionals?

- Degree of a vertex.
- Sum of the d-th powers of edge-lengths incident at a vertex. (d is the dimension)

Poisson point process

Definition

A point process $\mathcal{N}$ is a Poisson point process on $\mathbb{R}^d$ with intensity measure Λ if it satisfies the following:

Poisson point process

Definition

A point process $\mathcal{N}$ is a Poisson point process on $\mathbb{R}^d$ with intensity measure Λ if it satisfies the following:

- For any relatively compact set B (i.e. closure(B) is compact), the number of points of $\mathcal{N}$ in B, $|\mathcal{N} \cap B|$ is a Poisson random variable with parameter $\Lambda(B)$.

$$\mathbb{P}[|\mathcal{N} \cap B| = k] = e^{-\Lambda(B)} \frac{\Lambda(B)^k}{k!}$$

Poisson point process

Definition

A point process $\mathcal{N}$ is a Poisson point process on $\mathbb{R}^d$ with intensity measure Λ if it satisfies the following:

- For any relatively compact set B (i.e. closure(B) is compact), the number of points of $\mathcal{N}$ in B, $|\mathcal{N} \cap B|$ is a Poisson random variable with parameter $\Lambda(B)$.

$$\mathbb{P}[|\mathcal{N} \cap B| = k] = e^{-\Lambda(B)} \frac{\Lambda(B)^k}{k!}$$

- For any finite number of pairwise disjoint relatively compact sets $B_1, B_2, ..., B_n$, the random variables $|\mathcal{N} \cap B_1|, ..., |\mathcal{N} \cap B_n|$ are independent.

Stationary Poisson point process

If the Poisson point process $\mathcal{N}$ is stationary then

Stationary Poisson point process

If the Poisson point process $\mathcal{N}$ is stationary then

- The intensity measure $\Lambda(B) = \rho\lambda(B)$ (ρ is a constant, λ Lebesgue measure on $\mathbb{R}^d$).

Stationary Poisson point process

If the Poisson point process $\mathcal{N}$ is stationary then

- The intensity measure $\Lambda(B) = \rho\lambda(B)$ (ρ is a constant, λ Lebesgue measure on $\mathbb{R}^d$).
- Palm probability: Informally, it is the distribution of $\mathcal{N}$ conditioned that a typical point of this process is located at 0.

Stationary Poisson point process

If the Poisson point process $\mathcal{N}$ is stationary then

- The intensity measure $\Lambda(B) = \rho\lambda(B)$ (ρ is a constant, λ Lebesgue measure on $\mathbb{R}^d$).
- Palm probability: Informally, it is the distribution of $\mathcal{N}$ conditioned that a typical point of this process is located at 0.
- Its Palm probability is equal to $\mathcal{N} \cup \delta_0 := \mathcal{N}^o$.

Nice sets

Definition

A set $X = \{x_1, x_2, \dots\}$ of finite or countably finite points of $\mathbb{R}^d$ ($d \geq 2$) is called nice if

Nice sets

Definition

A set $X = \{x_1, x_2, ...\}$ of finite or countably finite points of $\mathbb{R}^d$ $(d \geq 2)$ is called nice if

- X is locally finite, i.e. every bounded subset of $\mathbb{R}^d$ has finite number of points of X.

Nice sets

Definition

A set $X = \{x_1, x_2, ...\}$ of finite or countably finite points of $\mathbb{R}^d$ $(d \geq 2)$ is called nice if

- X is locally finite, i.e. every bounded subset of $\mathbb{R}^d$ has finite number of points of X.
- All the interpoint distances are distinct, i.e. for any two distinct pair of points $\{x_1, x_2\} = \{x_3, x_4\}$, $|x_1 - x_2| \neq |x_3 - x_4|$.

Nice sets

Definition

A set $X = \{x_1, x_2, ...\}$ of finite or countably finite points of $\mathbb{R}^d$ ($d \geq 2$) is called nice if

- X is locally finite, i.e. every bounded subset of $\mathbb{R}^d$ has finite number of points of X.
- All the interpoint distances are distinct, i.e. for any two distinct pair of points $\{x_1, x_2\} = \{x_3, x_4\}$, $|x_1 - x_2| \neq |x_3 - x_4|$.

Examples

- The Poisson point process $\mathcal{N}$ is a.s. nice.

Nice sets

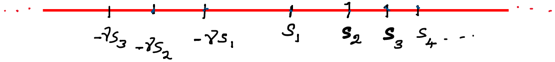
Definition

A set $X = \{x_1, x_2, ...\}$ of finite or countably finite points of $\mathbb{R}^d$ ($d \geq 2$) is called nice if

- X is locally finite, i.e. every bounded subset of $\mathbb{R}^d$ has finite number of points of X.
- All the interpoint distances are distinct, i.e. for any two distinct pair of points $\{x_1, x_2\} = \{x_3, x_4\}$, $|x_1 - x_2| \neq |x_3 - x_4|$.

Examples

- The Poisson point process $\mathcal{N}$ is a.s. nice.
- Let $s_n = \sum_{i=1}^{n} 1/i$ and γ be an irrational number. $X_n = \{-\gamma s_n, ..., -\gamma s_1, s_1, ..., s_n\}$



Construction of MST

Tree starting from a point

Let X be a nice set and $x \in X$. The tree starting from x, $t_\infty(x)$, is constructed inductively in the following way:

Construction of MST

Tree starting from a point

Let X be a nice set and $x \in X$. The tree starting from x, $t_\infty(x)$, is constructed inductively in the following way:

- Let $t_1(x)$ be the tree with vertex set $\{x\}$.

Construction of MST

Tree starting from a point

Let X be a nice set and $x \in X$. The tree starting from x, $t_\infty(x)$, is constructed inductively in the following way:

- Let $t_1(x)$ be the tree with vertex set $\{x\}$.
- Let $t_n(x)$ be the tree at obtained at nth step with vertex set $V_n = \{\eta_1, \eta_2, ...\eta_n\}$ (where $\eta_1 = x$). Let η_{n+1} be the point of $X \backslash V_n$ which is closest to the set V_n and $z \in V_n$ be such a closest point. Then, $t_{n+1}(x)$ is obtained by adding a new edge (η_{n+1}, z) to t_n.

Construction of MST

Tree starting from a point

Let X be a nice set and $x \in X$. The tree starting from x, $t_\infty(x)$, is constructed inductively in the following way:

- Let $t_1(x)$ be the tree with vertex set $\{x\}$.
- Let $t_n(x)$ be the tree at obtained at nth step with vertex set $V_n = \{\eta_1, \eta_2, ...\eta_n\}$ (where $\eta_1 = x$). Let η_{n+1} be the point of $X \backslash V_n$ which is closest to the set V_n and $z \in V_n$ be such a closest point. Then, $t_{n+1}(x)$ is obtained by adding a new edge (η_{n+1}, z) to t_n.
- $t_\infty(x) = \cup_{i=1}^{\infty} t_n(x)$.

Example

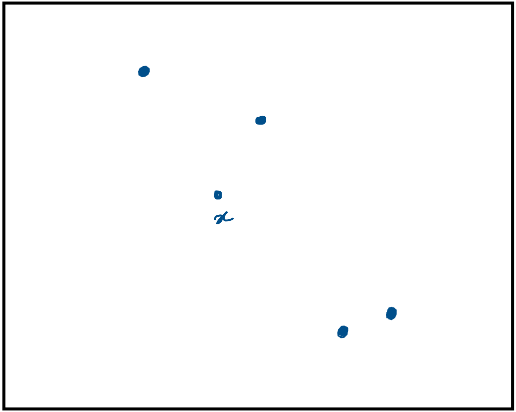


Figure: Construction of $t_n(x)$

Example

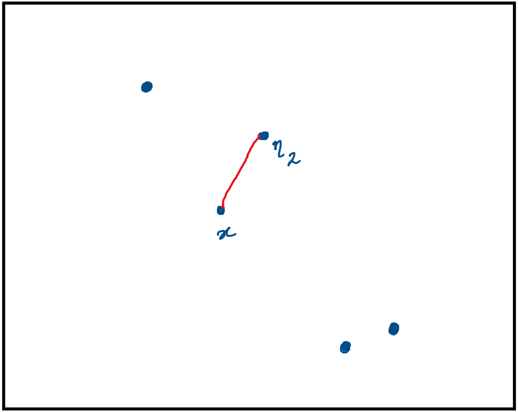


Figure: Construction of $t_n(x)$

Example

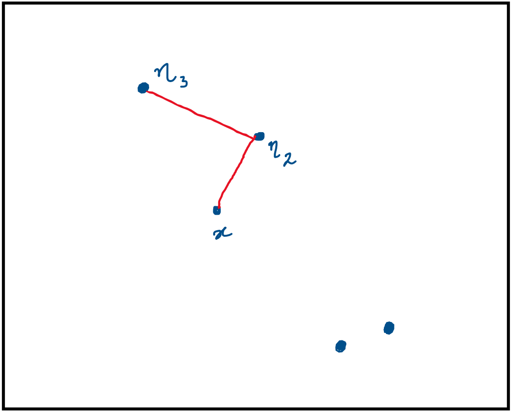


Figure: Construction of $t_n(x)$

Example

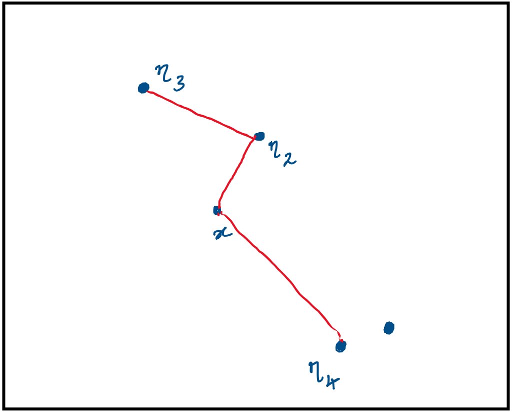


Figure: Construction of $t_n(x)$

Construction of MST

The graph $g(X)$

The graph $g(X)$ on a nice set X is defined in the following way:

Construction of MST

The graph $g(X)$

The graph $g(X)$ on a nice set X is defined in the following way:

- Vertex set of $g(X)$ is X.

Construction of MST

The graph $g(X)$

The graph $g(X)$ on a nice set X is defined in the following way:

- Vertex set of $g(X)$ is X.
- For $y_1, y_2 \in X$, an edge $(y_1, y_2) \in g(X)$ if $(y_1, y_2) \in t_\infty(y_1) \cup t_\infty(y_2)$.

Construction of MST

The graph $g(X)$

The graph $g(X)$ on a nice set X is defined in the following way:

- Vertex set of $g(X)$ is X.
- For $y_1, y_2 \in X$, an edge $(y_1, y_2) \in g(X)$ if $(y_1, y_2) \in t_\infty(y_1) \cup t_\infty(y_2)$.

Examples

- If $|X| = n$ then $g(X) = t_\infty(x) = t_n(x)$ for any $x \in X$.

Construction of MST

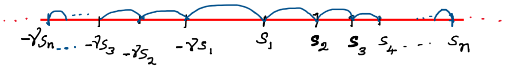
The graph $g(X)$

The graph $g(X)$ on a nice set X is defined in the following way:

- Vertex set of $g(X)$ is X.
- For $y_1, y_2 \in X$, an edge $(y_1, y_2) \in g(X)$ if $(y_1, y_2) \in t_\infty(y_1) \cup t_\infty(y_2)$.

Examples

- If $|X| = n$ then $g(X) = t_\infty(x) = t_n(x)$ for any $x \in X$.
- For $X_n = \{-\gamma s_n, ..., -\gamma s_1, s_1, ..., s_n\}$, $g(X_n) =$



Construction of MST

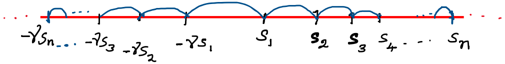
The graph $g(X)$

The graph $g(X)$ on a nice set X is defined in the following way:

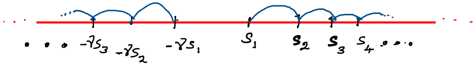
- Vertex set of $g(X)$ is X.
- For $y_1, y_2 \in X$, an edge $(y_1, y_2) \in g(X)$ if $(y_1, y_2) \in t_\infty(y_1) \cup t_\infty(y_2)$.

Examples

- If $|X| = n$ then $g(X) = t_\infty(x) = t_n(x)$ for any $x \in X$.
- For $X_n = \{-\gamma s_n, ..., -\gamma s_1, s_1, ..., s_n\}$, $g(X_n) =$



- For $X = \cup_{i \geq 1} X_n$, $g(X) =$



Properties of $g(X)$

Lemma

Let X be an infinite nice set. Then, $g(X)$ is a forest and the components are infinite.

Proof.

Properties of $g(X)$

Lemma

Let X be an infinite nice set. Then, $g(X)$ is a forest and the components are infinite.

Proof.

1. **Observation:** If an edge $(y_1, y_2) \in t_\infty(x)$ then either (y_1, y_2) is an edge of $t_\infty(y_1)$ or it is an edge of $t_\infty(y_2)$.

Properties of $g(X)$

Lemma

Let X be an infinite nice set. Then, $g(X)$ is a forest and the components are infinite.

Proof.

1. **Observation:** If an edge $(y_1, y_2) \in t_\infty(x)$ then either (y_1, y_2) is an edge of $t_\infty(y_1)$ or it is an edge of $t_\infty(y_2)$.
2. Let $x \in X$. Then by above observation $t_\infty(x) \subset g(X)$ proving that components are infinite.

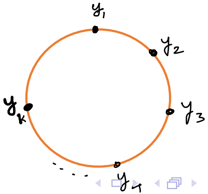
Properties of $g(X)$

Lemma

Let X be an infinite nice set. Then, $g(X)$ is a forest and the components are infinite.

Proof.

1. **Observation:** If an edge $(y_1, y_2) \in t_\infty(x)$ then either (y_1, y_2) is an edge of $t_\infty(y_1)$ or it is an edge of $t_\infty(y_2)$.
2. Let $x \in X$. Then by above observation $t_\infty(x) \subset g(X)$ proving that components are infinite.
3. Suppose there is a cycle $y_1, y_2, ...y_k, y_1$. Reorder them such that $|y_k - y_1|$ is the maximum among all distances. Then (y_1, y_k) is neither an edge $\in t_\infty(y_1)$ nor $\in t_\infty(y_k)$.



MSF of stationary Poisson point process, $g(\mathcal{N})$

Lemma

Let $\mathcal{N}^0 = \mathcal{N} \cup \delta_0$, $\mathcal{T}$ *be the connected component of* 0 *in* $g(\mathcal{N}^0)$, D *be the degree of* 0 *in* $\mathcal{T}$ *and* $L_1, L_2, ..., L_D$ *be the edge-lengths incident at* 0. *Then,*

1. $D \leq b_d$ *where* b_d *is a constant.*
2. $\mathbb{E}[D] = 2$.
3. $l_d = \sum_i \mathbb{E}[L_i^d] < \infty$.

Approximation of Poisson point process

Let $\mathcal{N}_n = \{\eta_1, \eta_2, ...\eta_n\}$ be i.i.d. points with uniform distribution on the unit cube $[0,1]^d$, $S_n = t_n(0, \mathcal{N}_n^*)$. Let $\mathcal{N}_n^* = \{n^{1/d}(\eta_i - \eta_1)\}$ be the scatter (scaled and shifted points) of $\mathcal{N}$. Then, we have the following proposition.

Approximation of Poisson point process

Let $\mathcal{N}_n = \{\eta_1, \eta_2, ...\eta_n\}$ be i.i.d. points with uniform distribution on the unit cube $[0,1]^d$, $S_n = t_n(0, \mathcal{N}_n^*)$. Let $\mathcal{N}_n^* = \{n^{1/d}(\eta_i - \eta_1)\}$ be the scatter (scaled and shifted points) of $\mathcal{N}$. Then, we have the following proposition.

Theorem

1. *$\mathcal{N}_n^* \to \mathcal{N}^0$ in distribution.*
2. *Let $\{e_i : i = 1, ...n-1\}$ be the edge-lengths of S_n. Then*

$$\sum_i^n |e_i|^d \to l_d \text{ in } L^2.$$

3. *Let $\Delta_{n,i}$ be the proportion of vertices of S_n with degree i, then for each i:*

$$\mathbb{E}[\Delta_{n,i}] \to \mathbb{P}[D = i].$$

Local convergence of finite sets

A sequence of nice sets X_n is said to converge locally to a nice set X is there exist a labelling of $X_n = \{x_{n1}, x_{n2}, ...\}$ and $X = \{x_1, x_2, ...\}$ such that:

1. $x_{ni} \to x_i$ for all i.
2. For any proper $C_L = [-L, L]^d$ (i.e., boundary of L does not intersect with X), $|X_n \cap C_L| \to |X \cap C_L|$.

Local convergence of graphs

Let $X_n = \{x_{n1}, x_{n2}, ...\}$ converge locally to $X = \{x_1, x_2, ...\}$, and h_n, h be graphs with vertex set X_n and X respectively, we say that h_n converge locally to h, if for any proper C_L, there exists $n_0 = n_0(L)$ such that for all $n \geq n_0$:

1. if (x_{ni}, x_{nj}) is an edge of h_n with $x_{ni} \in C_L$ then (x_i, x_j) is an edge of h.
2. if (x_i, x_j) is an edge of h with $x_i \in C_L$ then (x_{ni}, x_{nj}) is an edge of h_n.