

# Topological data analysis

## Homework 2

Sara Bizjak (27202020)

December 6, 2020

### 1 Theoretical problems

#### 1.1 Triangulations

a) Show that there are at most  $2^{\frac{n(n-1)}{2}}$  triangulations of  $S$ .

A triangulation is a decomposition of  $S$  into triangles, such that:

- no triangle is degenerate,
- interiors of the triangles are disjoint,
- the intersection of any two triangles is a common edge, a common vertex or empty.

In other words, a triangulation is a maximal crossing-free geometric graph on a point set  $S$ , or a subgraph of general plane graph (and also of maximal general plane graph).

A maximal general plane graph  $G$  (with vertices  $S$ ) is a graph with all possible connections between vertices and has  $(n-1) + (n-2) + \dots + 3 + 2 + 1$  connections, shortly written as  $\frac{n(n-1)}{2}$ .

If we take a look at subgraphs of  $G$ , we can quickly conclude that there are  $2^{\frac{n(n-1)}{2}}$  possible subgraphs, since with every connection we can choose to add it to the subgraph or not.

Since a triangulation is a subgraph of  $G$  and the number of all subsets of such graph is  $2^{\frac{n(n-1)}{2}}$ , we can conclude that there are at most  $2^{\frac{n(n-1)}{2}}$  triangulations of set  $S$ .

b) Constructing a set  $S$  such that all possible triangulations of  $S$  have a point of degree  $n-1$ . The degree of a point in a triangulation  $T$  is the number of edges in  $T$ , incident to that point. Therefore, in a set  $S$  there has to be a point that is connected to all other points, for example, the red point on a picture has always a degree  $n-1$ .

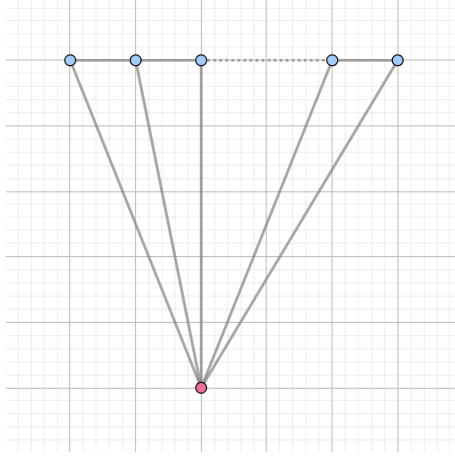


Figure 1: Construction of set  $S$ .

c) Show that any triangulation  $T$  of  $S$  has a point of degree 5 or less using the fact that if not all points in  $S$  are collinear, then any triangulation  $T$  of  $S$  has at most  $3n - 3$  edges ( $|E| \leq 3n - 3$ ). Proof with contradiction. Assume every point in  $T$  has a degree 6 or more, so  $\deg(v) \geq 6$  for all  $v \in V$ .

Since every connection has 2 edges, it follows that  $2|E| = |V| = \sum_{v \in V} \deg(v)$ .

By using the fact that all points in  $T$  has a degree at least 6, we can calculate that

$$2|E| = |V| = \sum_{v \in V} \deg(v) \geq \sum_{v \in V} 6 = 6n$$

and it follows that  $|E| \geq 3n$ . From the proposition we know that  $|E| \leq 3n - 3$  and if we combine both equations, we get

$$3n \leq |E| \leq 3n - 3 \Rightarrow 0 \leq -3$$

which is in contradiction, hence the proposal is true.

## 1.2 Vietoris-Rips Complex and Čech Complex

Let  $S = \{(0, 0), (2, 0), (1, 0.5), (1, 1.5)\} \in \mathbb{R}^2$ , where  $A = (0, 0)$ ,  $B = (2, 0)$ ,  $C = (1, 0.5)$ ,  $D = (1, 1.5)$ .

a) Building the Vietoris-Rips complex  $\text{Rips}(S, 2\epsilon)$  and the Čech complex  $\text{Cech}(S, \epsilon)$  for  $\epsilon = 0.8$ . In the Vietoris-Rips complex  $\text{Rips}(S, 2\epsilon)$  two vertices are connected by an edge if and only if the distance between them is less or equal to  $2\epsilon$ . To check that, we draw closed balls of radius  $\epsilon$  centered in points from given set  $S$ . We construct  $\text{Rips}(S, 2\epsilon)^{(1)}$  and the next thing we want to do is to find all possible complete subgraphs.

For  $2\epsilon = 1.6$  (or  $\epsilon = 0.8$ ) we get  $\text{Rips}(S, 1.6)^{(1)} = \{A, B, C, D, AC, BC, CD\}$ . There are no simplices of dimension 2 or higher in  $\text{Rips}(S, 1.6)^{(1)}$ , so  $\text{Rips}(S, 1.6) = \text{Rips}(S, 1.6)^{(1)}$  and  $\dim(\text{Rips}(S, 1.6)) = 1$  with Euler characteristic  $\chi = 4 - 3 = 1$ .

The Čech complex  $\text{Cech}(S, \epsilon)$  is the nerve of the cover of  $S$  with balls of radius  $\epsilon$ , so we can

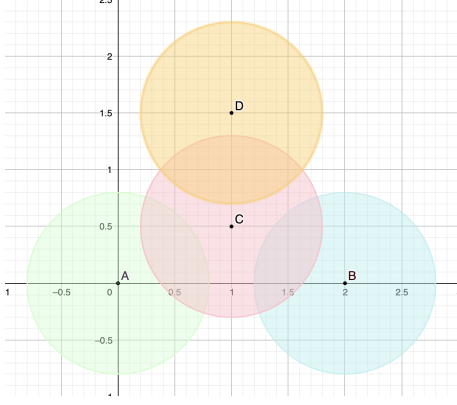


Figure 2:  $\epsilon = 0.8$ .

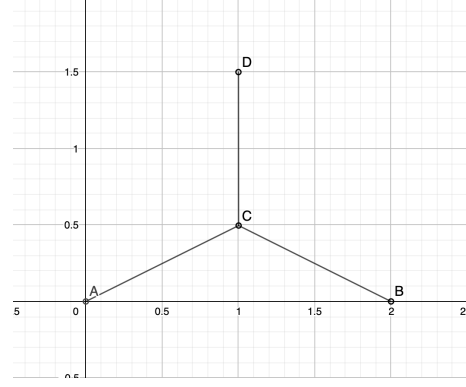


Figure 3:  $\epsilon = 0.8$ .

reuse the images. It is true that  $\text{Cech}(S, \epsilon)^{(1)} = \text{Rips}(S, 2\epsilon)^{(1)}$  and  $\text{Cech}(S, \epsilon) \subset \text{Rips}(S, 2\epsilon)$ . So the Vietoris-Rips and Čech complexes are the same in this case.

b) Building the Vietoris-Rips complex  $\text{Rips}(S, 2\epsilon)$  and the Čech complex  $\text{Cech}(S, \epsilon)$  for  $\epsilon = 1$ . For  $2\epsilon = 2$  (or  $\epsilon = 1$ ) we get  $\text{Rips}(S, 2)^{(1)} = \{A, B, C, D, AB, AC, AD, BC, BD, CD\}$ . We also have three copies of  $K_3$  ( $ABC, ACD, BCD$ ) and one of  $K_4$  ( $ABCD$ ), so  $\text{Rips}(S, 2) = \text{Rips}(S, 2)^{(1)} \cup \{ABC, ACD, BCD\} \cup \{ABCD\}$  and  $\dim(\text{Rips}(S, 2)) = 3$  with Euler characteristic  $\chi = 4 - 6 + 3 - 1 = 0$ .

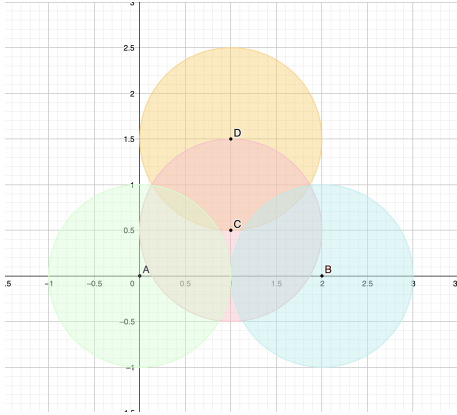


Figure 4:  $\epsilon = 1$ .

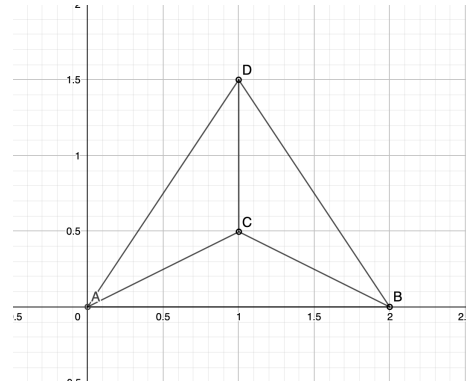


Figure 5:  $\epsilon = 1$ .

Since all of three circles centered in points  $A, B, C$ , points  $A, C, D$  and points  $B, C, D$  have a non-empty triple intersection (all four don't), all three 2-simplices from  $\text{Rips}(S, 2)$  are also contained in  $\text{Cech}(S, 1)$ .  $\text{Cech}(S, 1) = \{A, B, C, D, AB, AC, AD, BC, BD, CD, ABC, ACD, BCD\}$  and  $\dim(\text{Cech}(S, 1)) = 2$  with Euler characteristic  $\chi = 4 - 6 + 3 = 1$ .

### 1.3 Voronoi diagram

Finding a configuration of  $n$  points ( $n > 3$ ) in the plane such that their Voronoi diagram will have a cell with  $n - 1$  vertices.

The Voronoi diagram is a decomposition of  $\mathbb{R}^2$  into Voronoi regions, where the Voronoi region  $V_a$  of  $a \in S$  is  $V_a = \{x \in \mathbb{R}^2; d(x, a) \leq d(x, a') \forall a' \in S\}$ .

Such configuration of  $n$  points is  $n - 1$  points on the circle and one point in the middle of that circle.

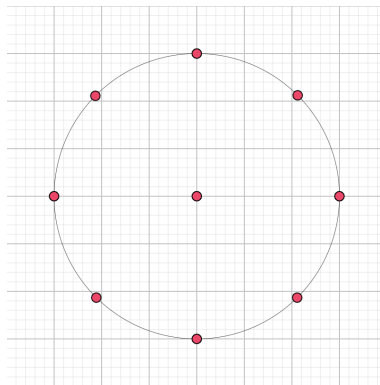


Figure 6: Example for  $n = 9$ .

### 1.4 Chessboard Complex

The chess board complex of a  $m \times n$  chessboard is a simplicial complex  $\Delta_{m \times n}$ . The vertices of  $\Delta_{m \times n}$  correspond to the squares of the chessboard. Simplices of  $\Delta_{m \times n}$  correspond to non-taking placements of rooks (ie. placements where no two rooks are in the same column or in the same row).

a) Showing that the chessboard complex  $\Delta_{3 \times 2}$  is homeomorphic to the circle  $S^1$ . (Showing that the simplicial complex you obtain is a 1-dimensional connected manifold with no boundary.)

If we follow the hint, we can read the proof of the picture. This is a 1-dimensional complex with a connected circular path with no boundary (all the vertices and edges are interior – two connections from one vertex).

$\Rightarrow$  Chessboard complex  $\Delta_{3 \times 2}$  is homeomorphic to the circle  $S^1$ .

b) Showing that the chessboard complex  $\Delta_{4 \times 3}$  is homeomorphic to a torus  $S^1 \times S^1$ . (Showing that the complex we obtain is an orientable connected 2-dimensional manifold without boundary with Euler characteristic 0.)

Following the hint:

If we take a look at the picture we can see that every point is connected to exactly 6 other points. That 6 points represent the 1-dim complex from task a), for which we know it is homeomorphic to the circle  $S^1$ . When we connect those 6 points to every point, we get 2-dimensional complex.

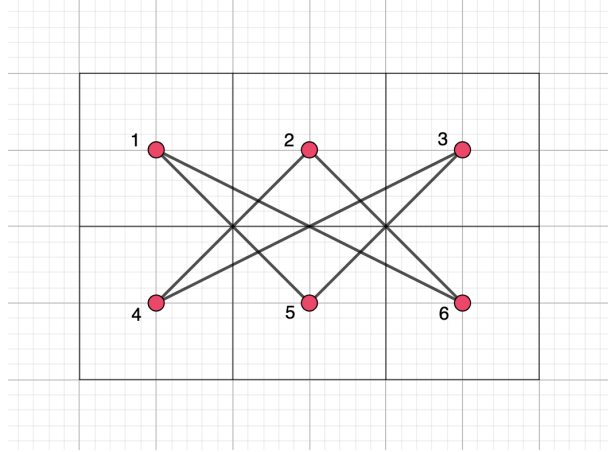


Figure 7:  $\Delta_{3 \times 2}$ .

This 2-dimensional complex has a connected circular path with no boundary, since each point has an even number of connections. The path is orientable – for a selected a point we can choose any other point and determin the orientation of surrounding of that point. We can also determin the orientation the other way around. In other words, we can "travel" around our manifold in any direction and are able to determin the orientation. This manifold has exactly two different possible orientations, so it is connected and orientable.

Euler characteristic is  $\chi = 12 - \frac{12 \cdot 6}{2} + \frac{2 \cdot \frac{6 \cdot 12}{2}}{3} = 12 - 36 + 24 = 0$ .

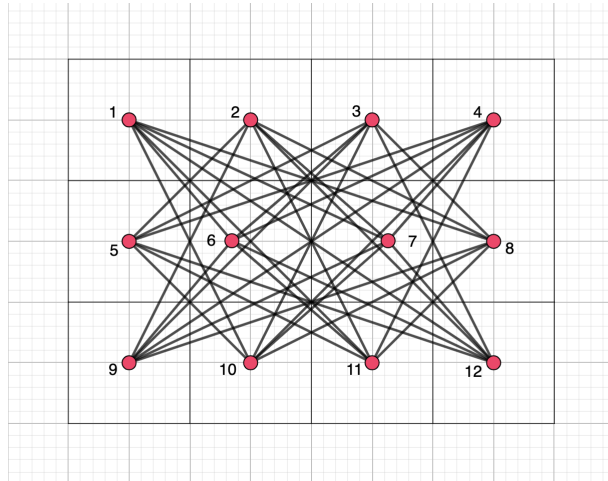


Figure 8:  $\Delta_{3 \times 2}$ .

$\Rightarrow$  Chessboard complex  $\Delta_{4 \times 3}$  is homeomorphic to a torus  $S^1 \times S^1$ .

c) Showing that the chessboard complex  $\Delta_{n \times (n-1)}$  is a manifold for all  $n$ .

From taskst a) and b) we know that  $\Delta_{3 \times 2}$  is a 1-dim manifold and  $\Delta_{4 \times 3}$  is a 2-dim manifold. With induction we would like to show that  $\Delta_{n \times (n-1)}$  is a  $(n - 2)$ -dim manifold assuming that  $\Delta_{(n-1) \times (n-2)}$  is a  $(n - 3)$ -dim manifold.

If we look at the pictures for complexes  $\Delta_{3 \times 2}$  and  $\Delta_{4 \times 3}$  we can see that we added a row and

a column with relevant connections to make a  $\Delta_{4 \times 3}$  from  $\Delta_{3 \times 2}$  manifold. So by adding new points and connections, we combined 1-dim manifold into 2-dim manifold.

Approaching the same idea we can combine  $(n-3)$ -dim manifold  $\Delta_{(n-1) \times (n-2)}$  and get  $\Delta_{n \times (n-1)}$  manifold (by adding  $n + n - 2$  points to  $\Delta_{(n-1) \times (n-2)}$  manifold – a row and a column), which is one dimension higher, so it is a  $(n-2)$ -dim manifold.

$\Rightarrow$  Chessboard complex  $\Delta_{n \times (n-1)}$  is a  $(n-2)$ -dim manifold for all  $n$ .

## 1.5 Vietoris-Rips Complex and Čech Complex in $d_\infty$ metric

Proving that if we define the Vietoris-Rips and the Čech Complex using  $d_\infty$  metric instead of  $d_2$ , then the complexes we obtain for any given  $\epsilon$  are the same, ie.  $C_\epsilon = VR_{2\epsilon}$ .

Metric  $d_\infty$ , also known as Chebyshev metric, is a metric defined on a vector space, where the distance between two vectors is the greatest of their differences along any coordinate dimensions.

$$d_\infty(x, y) := \max_i(|x_i - y_i|), (x_i, y_i) \in \mathbb{R}^d$$

First recall Vietoris-Rips and Čech complex definitions.

$$VR_{2\epsilon}(S) := \{T \subseteq S \mid \text{diam}(T) \leq \epsilon\}$$

$$C_\epsilon(S) := \{T \subseteq S \mid \bigcap_{x \in T} B_\epsilon(x) \neq \emptyset\}$$

To show  $C_\epsilon(S) = VR_{2\epsilon}(S)$  we will show the inclusions in both directions:  $C_\epsilon(S) \subseteq VR_{2\epsilon}(S)$  and  $VR_{2\epsilon}(S) \subseteq C_\epsilon(S)$ .

$C_\epsilon(S) \subseteq VR_{2\epsilon}(S)$ : The inclusion is obvious – if  $\epsilon$ -balls intersect then the diameter of their centers cannot exceed  $2\epsilon$ .

$VR_{2\epsilon}(S) \subseteq C_\epsilon(S)$ : Notice that under  $d_\infty$ -metric  $\epsilon$ -balls become hypercubes with side length  $2\epsilon$ . Two hypercubes intersect if and only if their centers are at most  $2\epsilon$  apart along all dimensions, that is

$$B_\epsilon(x) \cap B_\epsilon(y) \neq \emptyset \iff d(x_i, y_i) \leq 2\epsilon, \text{ where } (x_i, y_i) \in \mathbb{R}^d \iff d_\infty(x, y) \leq 2\epsilon$$

This generalizes to sets of balls, so that we have

$$\bigcap_{x \in T} B_\epsilon(x) \neq \emptyset \iff \text{diam}(T) \leq 2\epsilon$$

This proves the inclusion  $VR_{2\epsilon}(S) \subseteq C_\epsilon(S)$ .

$\Rightarrow C_\epsilon(S) = VR_{2\epsilon}(S)$  in  $d_\infty$  metric.

## 2 Programming problems

### 2.1 Line sweep triangulation

Let  $S$  be a cloud of points. We write a function `truangulate(S, vertical = True)` using a line sweep algorithm (by default the nime is vertical) that returns the list of edges  $E$  and triangles  $T$ .

All functions that we used are in the attached file `linesweeptriangulation.py`.

First, the algorithm sorts the points (whit small random amount of noise added) from left to right or from bottom up (depends on vertical parameter) and then go through all points in the determined direction. For current point we check to which (already checked point) we can connect the current one. That connection is realized if and only if it does not intersect with already made connections. When the algorithm checks this for all points in set  $S$ , it removes triangles that contain a point from the set  $S$ .

**Example 1:** Example from instructions.

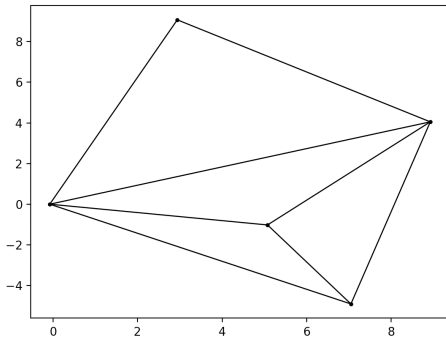


Figure 9: Horizontal.

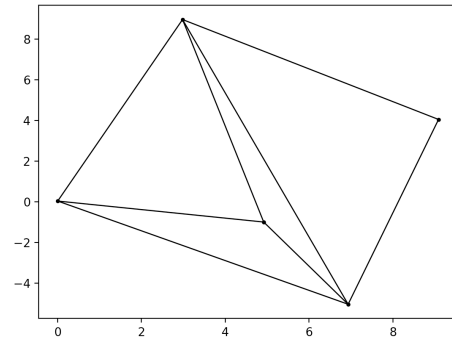


Figure 10: Vertical.

**Example 2:** First example on 100 random points.

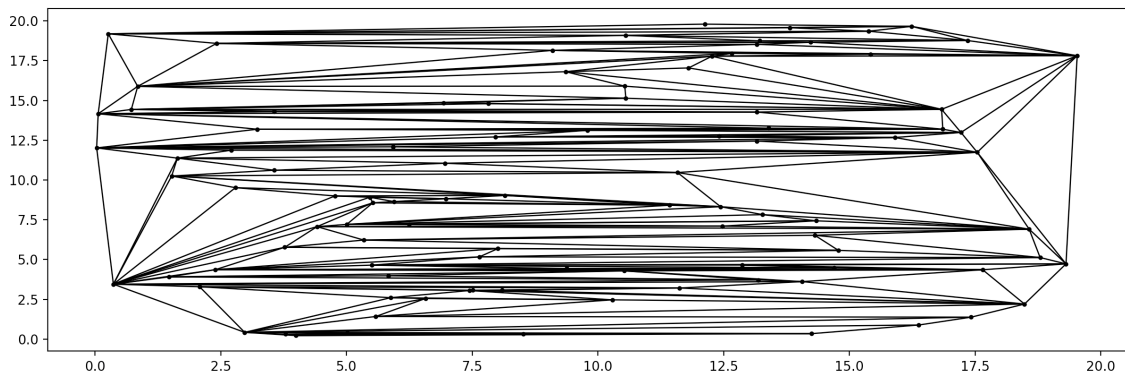


Figure 11: Horizontal.

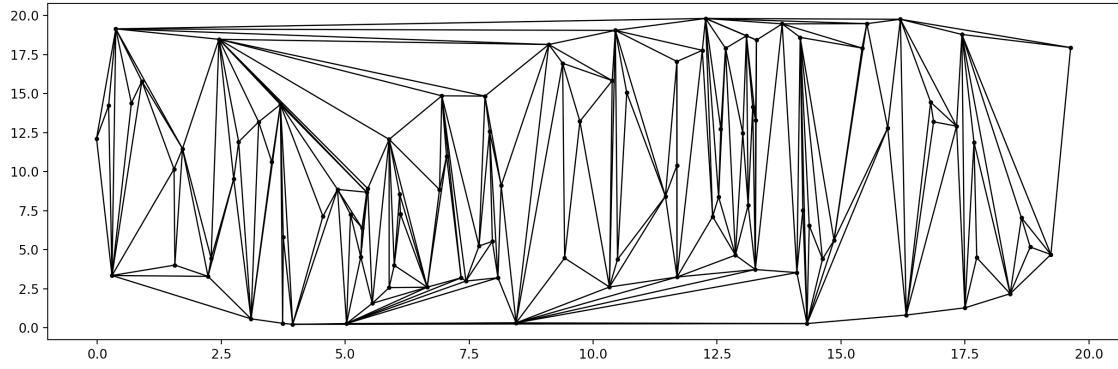


Figure 12: Vertical.

**Example 3:** Second example on 100 random points.

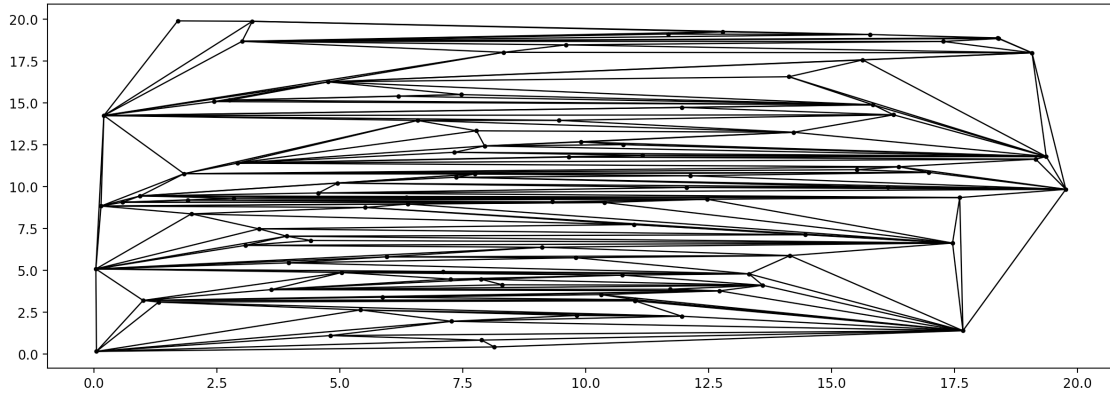


Figure 13: Horizontal.

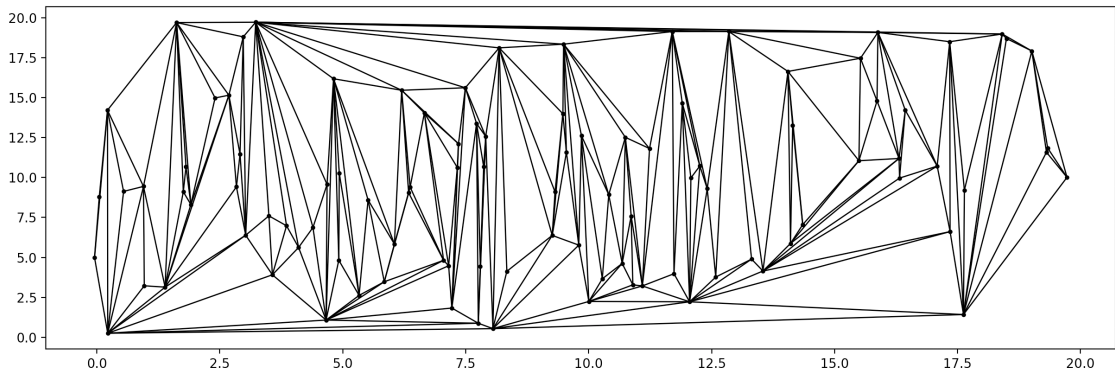


Figure 14: Vertical.



## 2.2 Delauney triangulation

Let  $T$  be any triangulation. A function `optimize(T)` optimizes  $T$  to produce the Delauney triangulation of the underlying set of points.

All functions that we used are in the attached file `delauney.py`.

First, the function makes a list of points in  $T$  and then by using python library `shapely` makes a Delauney triangulation.

**Example 1:** Example from instructions.

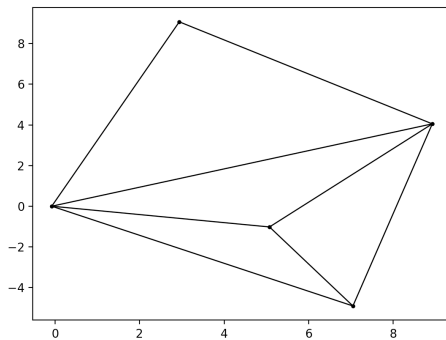


Figure 15: Triangulation.

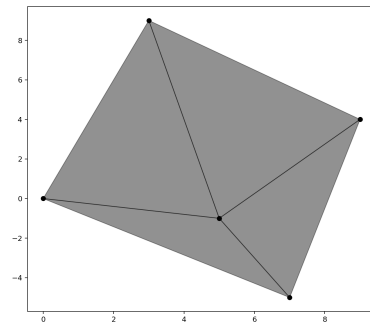


Figure 16: Delauney triangulation.

**Example 2:** First example on 100 random points.

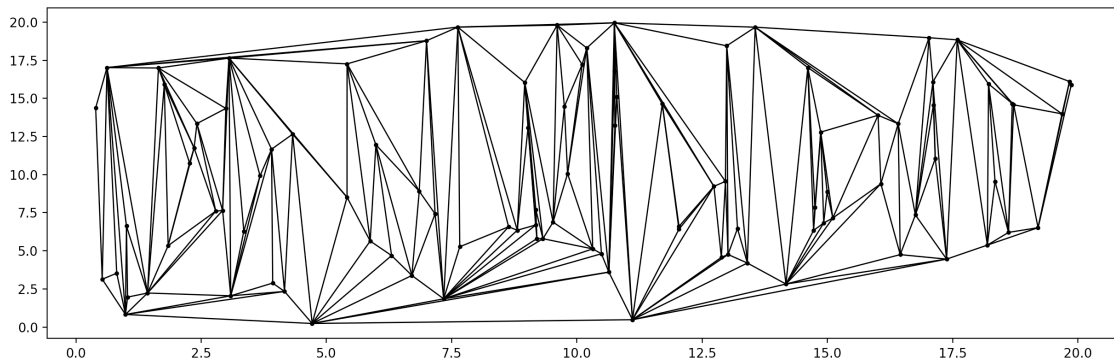


Figure 17: Triangulation.

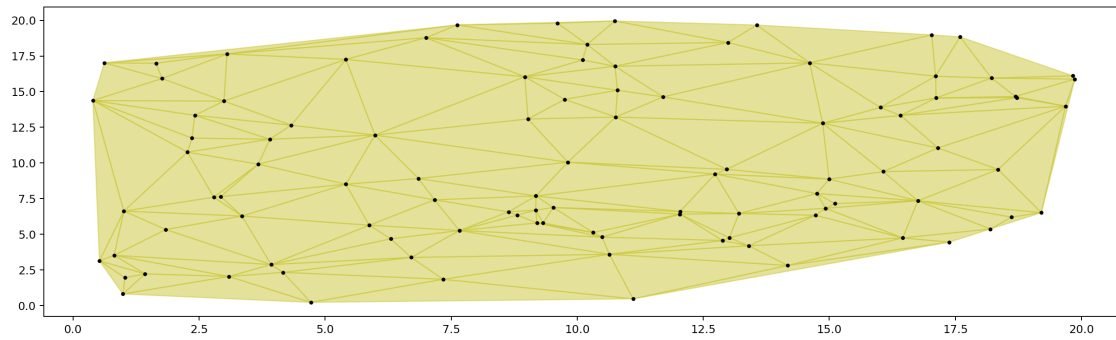


Figure 18: Delauney triangulation.

**Example 3:** Second example on 100 random points.

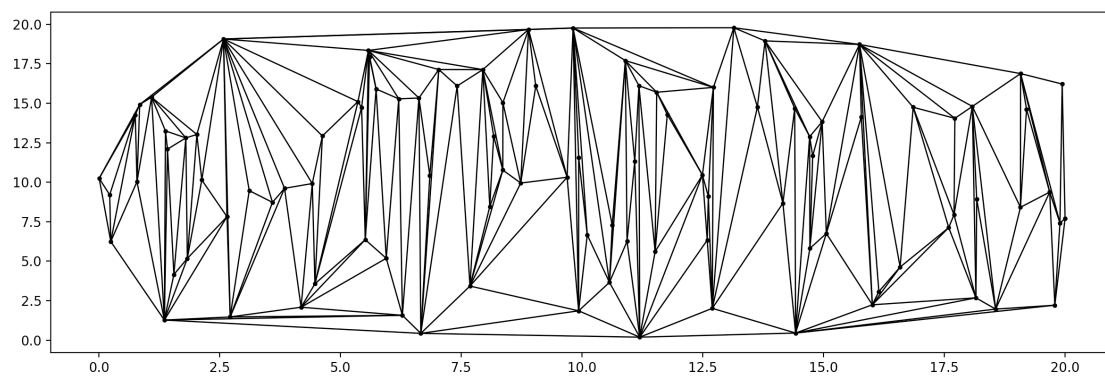


Figure 19: Triangulation.

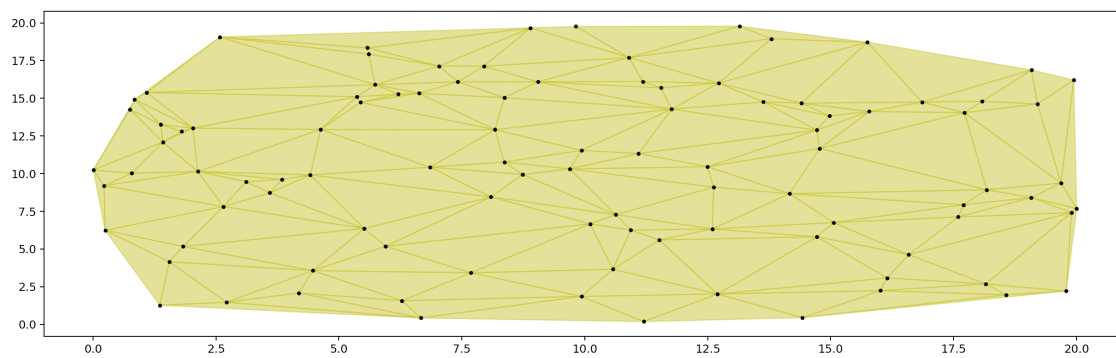


Figure 20: Delauney triangulation.

## 2.3 Orientation of surfaces