Chapter 5

Simplicial Complexes on Point Clouds

In general, the data we wish to analyze will not come in the form of a simplicial filtration, so in order to use persistent homology we need to transform our data into one. Ideally, the way we do this should retain the underlying shape of the data we want to analyze. In this section we discuss several ways of constructing simplicial complexes from point cloud data, and more generally, from finite metric spaces (i.e., a finite set of data points with given pairwise distances).

5.1 Čech and Vietoris-Rips complexes

Definition 5.1. *Given a metric space* (M, d)*, a finite point set* $P \subseteq M$*, and a real number radius* $r > 0$*, the* Čech complex $\mathbb{C}^r(P)$ *is defined as the nerve of the set of balls* $B(p, r) = \{x \in M \mid d(p, x) \leqslant r\}$ *for all* $p \in P$*.*

The Čech complex has the nice property that (at least for some metric spaces M including Euclidean space $\mathbb{R}^d$) by the Nerve theorem, it is homotopy equivalent to the union of the balls $B(p, r)$. In particular, for nice radii, it will capture the underlying shape. Sadly, checking whether a large number of balls have a common intersection can be computationally expensive. Further, the definition requires that the data points are embedded in a metric space. These two issues motivate the next definition.

Definition 5.2. *Given a finite metric space* (P, d) *and a real number radius* $r > 0$*, the* Vietoris-Rips complex $\mathbb{VR}^r(P)$ *is defined as the simplicial complex containing a simplex* σ *if and only if* $d(p, q) \leqslant 2r$ *for every pair* $p, q \in \sigma$*.*

Clearly, for finite subsets of metric spaces, by definition, the Čech complex and the Vietoris-Rips complex for the same radius and the same point set have the same set of 1-simplices (the same 1-skeleton). While the Čech complex then contains additional information about the common intersections of balls, the Vietoris-Rips complex is simply the clique complex of this 1-skeleton. This makes the Vietoris-Rips complex easier to compute. Furthermore, we make the following simple observation, showing that the Vietoris-Rips complex still approximately captures shapes in the data:

Observation 5.3. $\mathbb{C}^r(P) \subseteq \mathbb{VR}^r(P) \subseteq \mathbb{C}^{2r}(P)$.

Exercise 5.4. *Prove Observation 5.3.*

Exercise 5.5. *Find a point set* $P \subset \mathbb{R}^2$ *and a radius* r *such that its Vietoris-Rips complex has non-trivial 2-homology, i.e., such that* $H_2(\mathbb{VR}^r(P)) \not\cong 0$.
Furthermore, is there a dimension k *such that* $H_{k'}(\mathbb{VR}^r(Q)) = 0$ *for all* $k' \geqslant k$*, all* $r > 0$*, and all point sets* $Q \subset \mathbb{R}^2$*?*

5.2 Delaunay and Alpha complexes

Recall that computing persistent homology takes $O(N^3)$ time, where N is the size of the simplicial complex in the filtration. For large enough radii, both the Čech and the Vietoris-Rips complex become complete, and thus contain 2^n simplices. Computing persistent homology using those complexes is therefore computationally very expensive, which is why in many applications we would like to have sparser complexes. For data in $\mathbb{R}^d$ we can look at the so-called Delaunay triangulation, which only has complexity $O(n^{\lceil d/2 \rceil})$.

Definition 5.6. *Given a finite point set* $P \subset \mathbb{R}^d$*, a* Delaunay simplex *is a geometric simplex whose vertices are in* P *and lie on the boundary of a ball whose interior contains no points of* P.

A Delaunay triangulation $\mathrm{Del}(P)$ *of* P *is a geometric simplicial complex with the vertex set* P *where every simplex is a Delaunay simplex and whose underlying space covers the convex hull of* P.

Given a finite point set $P \subset \mathbb{R}^d$*, the* extended Delaunay complex *is the simplicial complex where for every face* σ*, for* $d' \leqslant d$*, every* d'*-face of* σ *is a Delaunay simplex.*

It is a well-known fact that for a point set in general position (no $d+2$ points lie on a common sphere), there is a unique Delaunay triangulation. Furthermore, in this case the extended Delaunay complex and this unique Delaunay triangulation coincide.

Definition 5.7. *Given a finite point set* $P \subset \mathbb{R}^d$*, the* Voronoi diagram *is the tessellation of* $\mathbb{R}^d$ *into the* Voronoi cells

$$V_p = \{x \in \mathbb{R}^d \mid d(x,p) \leqslant d(x,q) \forall q \in P\}$$

for all $p \in P$.

Fact 5.8. *The nerve of the Voronoi cells of* P *is the extended Delaunay complex of* P.

Exercise 5.9. *Convince yourself that for a point set in* $\mathbb{R}^2$*, the nerve of the Voronoi diagram is the extended Delaunay complex. Furthermore, convince yourself that if the points are in general position (there are no three points that are collinear, and no four points that are cocircular), then there is a unique Delaunay triangulation.*

Based on the Delaunay triangulation, we define the *Alpha complex* by parameterizing using a radius as follows:

Definition 5.10. *Given a finite point set* $P \subset \mathbb{R}^d$ *in general position as well as a real number radius* $r > 0$*, the* Alpha complex $\mathrm{Del}^r(P)$ *consists of all simplices* $\sigma \in \mathrm{Del}(P)$ *for which the circumscribing ball of* σ *has radius at most* r*.*

The following fact provides us with an alternative definition of the Alpha complex:

Fact 5.11. *The Alpha complex* $\mathrm{Del}^r(P)$ *is the nerve of the sets* $B(p, r) \cap V_p$ *for all* $p \in P$*.*

Since the Alpha complex is a subset of the Delaunay triangulation (and for large enough radius is equal to the Delaunay triangulation), it also has complexity $O(n^{\lceil d/2 \rceil})$. Further, the above fact together with the Nerve theorem implies that the Alpha complex $\mathrm{Del}^r(P)$ is homotopy equivalent to the Čech complex $\mathbb{C}^r(P)$.

Exercise 5.12. *Is the following true or false? Consider a point set* $P \subset \mathbb{R}^2$ *in general position and a radius* $r > 0$*. Then the Alpha complex (with radius* r*) is the intersection of the Čech complex (with radius* r*) with the Delaunay triangulation.*

5.3 Subsample Complexes

For many applications, the Alpha complex is still too large. It is further expensive to compute, as computing a Delaunay triangulation in $\mathbb{R}^d$ takes $O(n^{\lceil d/2 \rceil})$ time. Sparser complexes can be constructed by looking at subsamples of the data, and relating the rest of the data to these subsamples. In the following, we will discuss two examples of complexes based on this idea.

Definition 5.13. *Given a finite point set* Q *and a point set* $P \supset Q$ *in some metric space, we say that a simplex* $\sigma \subseteq Q$ *is* weakly witnessed *by* $x \in P \setminus Q$*, if* $d(q, x) \leqslant d(p, x)$ *for every* $q \in \sigma$ *and* $p \in Q \setminus \sigma$*.*

Note that the set of weakly witnessed simplices is not downwards closed. We thus define a simplicial complex by requiring that all faces are weakly witnessed:

Definition 5.14. *The* Witness complex $\mathbb{W}(Q, P)$ *is the collection of simplices on* Q *for which every face is weakly witnessed by some point in* $P \setminus Q$*.*

Note that if we take the metric space $\mathbb{R}^d$ and we let P be the whole $\mathbb{R}^d$, then $\mathbb{W}(Q, P) = \mathrm{Del}(Q)$, and by definition we thus get in general that $\mathbb{W}(Q, P) \subseteq \mathrm{Del}(Q)$.

To arrive at a filtration, we again have to introduce a parameter $r > 0$:

Definition 5.15. *Given a finite point set* Q *and a point set* $P \supset Q$ *in some metric space as well as a real number radius* $r > 0$*, the* parameterized *Witness complex* $\mathbb{W}^r(Q, P)$ *is defined as follows:*
An edge pq *is in* $\mathbb{W}^r(Q, P)$ *if it is weakly witnessed by* $x \in P \setminus Q$ *and* $d(p, x) \leqslant r$ *and* $d(q, x) \leqslant r$*. A simplex* σ *is in* $\mathbb{W}^r(Q, P)$ *if all its edges are.*

Note that from this definition it is not guaranteed that the parameterized Witness complex is a subcomplex of the Witness complex.

The idea of the parameterized Witness complex is that it should approximate the Vietoris-Rips complex on P. There are theoretical guarantees about this approximation for manifolds of dimension at most 2, but the parameterized witness complex may fail to capture the topology of manifolds in dimension 3 and above.

Let us now consider a second subsample complex, the graph induced complex.

Definition 5.16. *Given two finite point sets* Q, P *in* $\mathbb{R}^d$*, as well as a graph* $G(P)$ *with vertices in* P*, we define* $\nu : P \to Q$ *by sending each point in* P *to its closest point in* Q*. The* graph induced complex $\mathbb{G}(Q, G(P))$ *contains a simplex* $\sigma = \{q_0, \dots, q_k\} \subset Q$ *if and only if there is a clique* $\{p_0, \dots, p_k\}$ *in* $G(P)$ *for which* $\nu(p_i) = q_i$.

We again parameterize this:

Definition 5.17. *Let* $G^r(P)$ *be the graph on* P *where* pq *is an edge if and only if* $d(p, q) \leqslant 2r$*. The* parameterized graph induced complex $\mathbb{G}^r(Q, P)$ *is defined as* $\mathbb{G}(Q, G^r(P))$.

This complex again has theoretical guarantees of approximating the Vietoris-Rips complex on $P \cup Q$.

Exercise 5.18. *Let* P, Q *be point sets and* $G(P)$ *a graph with* P *as its vertex set. Let* $\nu : P \to Q$ *be the map sending each point of* P *to its closest point of* Q *(assume that this closest point is always unique). Let* C *be the clique complex of* $G(P)$ *(the complex which includes a simplex iff its corresponding vertices in* $G(P)$ *form a clique).*

Show that ν *extends to a simplicial map* $\bar{\nu} : C \to \mathbb{G}(Q, G(P))$*. Also show that any simplicial complex* K *with* $V(K) = Q$ *for which* ν *has a simplicial extension must contain* $\mathbb{G}(Q, G(P))$.

Questions

16. *What are the Čech and Vietoris-Rips complexes?* Give the definitions, discuss their size and theoretical guarantees, and how they are related.
17. *What are the Delaunay and Alpha complexes?* Give the definitions, discuss their size and theoretical guarantees, and how they are related.
18. *What is the Witness complex?* State the Definition and describe how it relates to the non-sparse complexes.
19. *What is the Graph induced complex?* State the Definition and describe how it relates to the non-sparse complexes.