A DISTANCE BETWEEN FILTERED SPACES VIA TRIPODS

FACUNDO MÉMOLI

Abstract. We present a simplified treatment of stability of filtrations on finite spaces. Interestingly, we can lift the stability result for combinatorial filtrations from [CSEM06] to the case when two filtrations live on different spaces without directly invoking the concept of interleaving.

Contents

1. Introduction
2. Simplicial Homology
3. Filtrations and Persistent Homology
4. Stability of filtrations
   4.1. Filtrations defined over different sets
       4.1.1. Common parametrizations of two spaces: tripods.
5. Filtrations arising from metric spaces: Rips and Čech
   5.1. The Rips filtration
   5.2. The Čech filtration
6. Discussion
References

1. Introduction

My goal for the construction that I describe in this note was to lift the stability result of [CSEM06] to the setting when the simplicial filtrations are not necessarily defined on the same set. The ideas in this note were first presented at ATCMS in July 2012. A partial discussion appears in [Ban12].

2. Simplicial Homology

Given a simplicial complex $L$ and simplices $\sigma, \tau \in L$, we write $\sigma \subseteq \tau$ whenever $\sigma$ is a face of $\tau$. For each integer $\ell \geq 0$ we denote by $L^{(\ell)}$ the $\ell$-skeleton of $L$.

Recall that given two finite simplicial complexes $L$ and $S$, a simplicial map between them arises from any map $f : L^{(0)} \to S^{(0)}$ with the property that whenever $p_0, p_1, \ldots, p_k$ span a simplex in $L$, then $f(p_0), f(p_1), \ldots, f(p_k)$ span a simplex of $S$. One does not require that the vertices $f(p_0), f(p_1), \ldots, f(p_k)$ be all distinct. Given a map $f : L^{(0)} \to S^{(0)}$ between the vertex sets of the finite simplicial complexes $L$ and $S$, we let $\overline{f} : L \to S$ denote the induced simplicial map.

Date: April 14, 2017.

This work was supported by NSF grants IIS-1422400 and CCF-1526513.
We will make use of the following theorem in the sequel.

**Theorem 2.1** (Quillen’s Theorem A in the simplicial category, [Qui73]). Let \( \zeta : S \to L \) be a simplicial map between two finite complexes. Suppose that the preimage of each closed simplex of \( L \) is contractible. Then \( \zeta \) is a homotopy equivalence.

**Corollary 2.1.** Let \( L \) be a finite simplicial complex and \( \varphi : Z \to L^{(0)} \) be any surjective map with finite domain \( Z \). Let \( S := \{ \tau \subseteq Z | \varphi(\tau) \in L \} \). Then \( S \) is a simplicial complex and the induced simplicial map \( \varphi : S \to L \) is an homotopy equivalence.

**Proof.** Note that \( S = \bigcup_{\sigma \in L} \{ \tau \subseteq Z | \varphi(\tau) = \sigma \} \) so it is clear that \( S \) is a simplicial complex with vertex set \( Z \). That the preimage of each \( \sigma \in L \) is contractible is trivially true since those preimages are exactly the simplices in \( S \). The conclusion follows directly from Quillen’s Theorem A.

In this paper we consider homology with coefficients in a field \( \mathbb{F} \) so that given a simplicial complex \( L \), then for each \( k \in \mathbb{N} \), \( H_k(L, \mathbb{F}) \) is a vector space. To simplify notation, we drop the argument \( \mathbb{F} \) from the list and only write \( H_k(L) \) for the homology of \( L \) with coefficients in \( \mathbb{F} \).

### 3. Filtrations and Persistent Homology

Let \( \mathcal{F} \) denote the set of all finite filtered spaces: that is pairs \( X = (X, F_X) \) where \( X \) is a finite set and \( F_X : \text{pow}(X) \to \mathbb{R} \) is a monotone function. Any such function is called a *filtration* over \( X \). Monotonicity in this context refers to the condition that \( F_X(\sigma) \geq F_X(\tau) \) whenever \( \sigma \supseteq \tau \). Given a finite set \( X \), by \( F(X) \) we denote the set of all possible filtrations \( F_X : \text{pow}(X) \to \mathbb{R} \) on \( X \). Given a filtered space \( X = (X, F_X) \in \mathcal{F} \), for each \( \varepsilon \in \mathbb{R} \) define the simplicial complex

\[
L_{\varepsilon}(X) := \{ \sigma \subseteq X | F_X(\sigma) \leq \varepsilon \}.
\]

One then considers the nested family of simplicial complexes

\[
L(X) := \{ L_{\varepsilon}(X) \subset L_{\varepsilon'}(X) \}_{\varepsilon \leq \varepsilon'}
\]

where each \( L_{\varepsilon}(X) \) is, by construction, finite. At the level of homology, for each \( k \in \mathbb{N} \) the above inclusions give rise to a system of vector spaces and linear maps

\[
V_k(X) := \{ V_{\varepsilon}(X) \xrightarrow{\nu_{\varepsilon, \varepsilon'}} V_{\varepsilon'}(X) \}_{\varepsilon \leq \varepsilon'},
\]

which is called a *persistence vector space*. Note that each \( V_{\varepsilon}(X) \) is finite dimensional.

Persistence vector spaces admit a *classification up to isomorphism* in terms of collections of intervals so that to the persistence vector space \( V \) one assigns a multiset of intervals \( I(V) \) [CDS10]. These collections of intervals are sometimes referred to as *barcodes* or also *persistence diagrams*, depending on the graphical representation that is adopted [EH10]. We denote by \( \mathcal{D} \) the collection of all finite persistence diagrams. An element \( D \in \mathcal{D} \) is a finite multiset of points

\[
D = \{ (b_\alpha, d_\alpha), \ 0 \leq b_\alpha \leq d_\alpha, \ \alpha \in A \}
\]

for some (finite) index set \( A \). Given \( k \in \mathbb{N} \), to any filtered set \( X \in \mathcal{F} \) one can attach a persistence diagram via

\[
X \mapsto L(X) \mapsto V_k(X) \mapsto I(V_k(X)).
\]
We denote by \( D_k : \mathcal{F} \to \mathcal{D} \) the resulting composite map. Given \( X = (X, F_X) \), we will sometimes write \( D_k(F_X) \) to denote \( D_k(X) \).

4. Stability of filtrations

The bottleneck distance is a useful notion of distance between persistence diagrams and we recall it’s definition next. We will follow the presentation on \([Car14]\). Let \( \Delta \subset \mathbb{R}^2_+ \) be comprised of those points which sit above the diagonal: \( \Delta := \{(x, y) | x \leq y \} \).

Define the persistence of a point \( P = (x_P, y_P) \in \Delta \) by \( \text{pers}(P) := y_P - x_P \).

Let \( D_1 = \{P_\alpha\}_{\alpha \in A_1} \) and \( D_2 = \{Q_\alpha\}_{\alpha \in A_2} \) be two persistence diagrams indexed over the finite index sets \( A_1 \) and \( A_2 \), respectively. Consider subsets \( B_i \subseteq A_i \) with \( |B_1| = |B_2| \) and any bijection \( \varphi : B_1 \to B_2 \), then define

\[
J(\varphi) := \max \left( \max_{\beta \in B_1} \|P_\beta - Q_{\varphi(\beta)}\|_{\infty}, \max_{\alpha \in A_1 \setminus B_1} \frac{1}{2} \text{pers}(P_\alpha), \max_{\alpha \in A_2 \setminus B_2} \frac{1}{2} \text{pers}(P_\alpha) \right).
\]

Finally, one defines the bottleneck distance between \( D_1 \) and \( D_2 \) by

\[
d_\Delta(D_1, D_2) := \min_{(B_1, B_2, \varphi)} J(\varphi),
\]

where \((B_1, B_2, \varphi)\) ranges over all \( B_1 \subset A_1, \ B_2 \subset A_2 \), and bijections \( \varphi : B_1 \to B_2 \).

One of the standard results about the stability of persistent homology invariants, which is formulated in terms of the Bottleneck distance, is the proposition below which we state in a weaker form that will suffice for our presentation.

**Theorem 4.1** \([CSEM06]\). For all finite sets \( X \) and filtrations \( F, G : \text{pow}(X) \to \mathbb{R} \),

\[
d_\Delta(D_k(F), D_k(G)) \leq \max_{\sigma \in \text{pow}(X)} |F(\sigma) - G(\sigma)|,
\]

for all \( k \in \mathbb{N} \).

The proof of this theorem offered in \([CSEM06]\) is purely combinatorial and elementary. This result requires that the two filtrations be given on the same set. This restriction will be lifted using the ideas that follow.

4.1. Filtrations defined over different sets. A parametrization of a finite set \( X \) is any finite set \( Z \) and a surjective map \( \varphi_X : Z \to X \). Consider a filtered space \( X = (X, F_X) \in \mathcal{F} \) and a parametrization \( \varphi_X : Z \to X \) of \( X \). By \( \varphi_X^* F_X \) we denote the pullback filtration induced by \( F_X \) and the map \( \varphi_X \) on \( Z \). This filtration is given by \( \tau \mapsto F_X(\varphi_X(\tau)) \) for all \( \tau \in \text{pow}(Z) \).

A useful corollary of the persistence homology isomorphism theorem \([EHT10]\) pp. 139] and Corollary 2.1 is that the persistence diagrams of the original filtration and the pullback filtration are identical.

**Corollary 4.1.** Let \( X = (X, F_X) \in \mathcal{F} \) and \( \varphi : Z \to X \) a parametrization of \( X \). Then, for all \( k \in \mathbb{N} \), \( D_k(\varphi^* F_X) = D_k(F_X) \).

\[\text{In} \ [CSEM06] \text{the authors do not assume that the underlying simplicial complex is the full powerset.}\]
4.1.1. **Common parametrizations of two spaces: tripods.** Now, given $X = (X, F_X)$ and $Y = (Y, F_Y)$ in $\mathcal{F}$, the main idea in comparing filtrations defined on different spaces is to consider parametrizations $\varphi_X : Z \to X$ and $\varphi_Y : Z \to Y$ of $X$ and $Y$ from a common parameter space $Z$, i.e. **tripods**:

$$
\begin{array}{ccc}
Z & \xrightarrow{\varphi_X} & X \\
\downarrow & & \downarrow \\
Y & \xrightarrow{\varphi_Y} & 
\end{array}
$$

and compare the pullback filtrations $\varphi_X^* F_X$ and $\varphi_Y^* F_Y$ on $Z$. Formally, define

(1) \[ d_{\mathcal{F}}(X, Y) := \]
\[ \inf \left\{ \max_{\tau \in \text{pow}(Z)} |\varphi_X^* F_X(\tau) - \varphi_Y^* F_Y(\tau)| ; \varphi_X : Z \to X, \varphi_Y : Z \to Y \text{ parametrizations} \right\}. \]

**Remark 4.1.** Notice that in case $X = \{ \ast \}$ and $F_{\{\ast\}}(\ast) = c \in \mathbb{R}$, then $d_{\mathcal{F}}(X, Y) = \max_{\sigma \subseteq Y} |F_Y(\sigma) - c|$, for any filtered space $Y$. If $c = 0, Y = \{ y_1, y_2 \}$ with $F_Y(y_1) = F_Y(y_2) = 0$ and $F_Y(\{y_1, y_2\}) = 1$. Then, $d_{\mathcal{F}}(X, Y) = 1$.

However, still with $c = 0$ and $Y = \{ y_1, y_2 \}$, but $F_Y(y_1) = F_Y(y_2) = F_Y(\{ y_1, y_2 \}) = 0$, one has $d_{\mathcal{F}}(X, Y) = 0$. This means that $d_{\mathcal{F}}$ is at best a pseudometric on filtered spaces.

**Proposition 4.1.** $d_{\mathcal{F}}$ is a pseudometric on $\mathcal{F}$.

**Proof.** Symmetry and non-negativity are clear. We need to prove the triangle inequality. Let $X = (X, F_X), Y = (Y, F_Y), W = (W, F_W)$ in $\mathcal{F}$ be non-empty and $\eta_1, \eta_2 > 0$ be s.t.

$$
\begin{align*}
&d_{\mathcal{F}}(X, Y) < \eta_1 \\
&d_{\mathcal{F}}(Y, W) < \eta_2.
\end{align*}
$$

Choose, $\psi_X : Z_1 \to X, \psi_Y : Z_1 \to Y, \zeta_Y : Z_2 \to Y, \text{ and } \zeta_W : Z_2 \to W$ surjective such that

\[ \|F_X \circ \psi_X - F_Y \circ \psi_Y\|_{\ell^\infty(\text{pow}(Z_1))} < \eta_1 \]

and

\[ \|F_Y \circ \zeta_Y - F_W \circ \zeta_W\|_{\ell^\infty(\text{pow}(Z_2))} < \eta_2. \]

Let $Z \subseteq Z_1 \times Z_2$ be defined by $Z := \{(z_1, z_2) \in Z_1 \times Z_2 | \psi_Y(z_1) = \zeta_Y(z_2)\}$ and consider the following (pullback) diagram:

$$
\begin{array}{ccc}
Z & \xrightarrow{\pi_1} & Z_1 \\
\downarrow & & \downarrow \\
Y & \xrightarrow{\psi_Y} & W \\
\downarrow & & \downarrow \\
X & \xrightarrow{\psi_X} & Z_2 \\
\downarrow & & \downarrow \\
W & \xrightarrow{\zeta_W} & \n
\end{array}
$$

Clearly, since $\psi_Y$ and $\zeta_Y$ are surjective, $Z$ is non-empty. Now, consider the following three maps with domain $Z$: $\phi_X := \psi_X \circ \pi_1, \phi_Y := \psi_Y \circ \pi_1 = \zeta_Y \circ \pi_2$, and $\phi_W := \zeta_W \circ \pi_2$. These three maps...
maps are surjective and therefore constitute parametrizations of $X$, $Y$, and $W$, respectively. Then, since $\pi_i: Z \to Z_i$, $i = 1, 2$, are surjective and $\psi_Y \circ \pi_1 = \zeta_Y \circ \pi_2$, we have
\[
d_F(X, W) \leq \|F_X \circ \phi_X - F_W \circ \phi_W\|_{\ell^\infty(\text{pow}(Z))}
\leq \|F_X \circ \phi_X - F_Y \circ \phi_Y\|_{\ell^\infty(\text{pow}(Z))} + \|F_Y \circ \phi_Y - F_W \circ \phi_W\|_{\ell^\infty(\text{pow}(Z))}
= \|F_X \circ \psi_X - F_Y \circ \psi_Y\|_{\ell^\infty(\text{pow}(Z_1))} + \|F_Y \circ \zeta_Y - F_W \circ \zeta_W\|_{\ell^\infty(\text{pow}(Z_2))}
\leq \eta_1 + \eta_2.
\]
The conclusion follows by letting $\eta_1 \searrow d_F(X, Y)$ and $\eta_2 \searrow d_F(Y, W)$. \hfill \Box

We now obtain a lifted version of Theorem 4.1 below.

**Theorem 4.2.** For all finite filtered spaces $X = (X, F_X)$ and $Y = (Y, F_Y)$, and all $k \in \mathbb{N}$ one has:
\[
d_D(D_k(X), D_k(Y)) \leq d_F(X, Y).
\]

**Proof of Theorem 4.2.** Assume $\varepsilon > 0$ is such that $d_F(F_X, F_Y) < \varepsilon$. Then, let $\varphi_X: Z \to X$ and $\varphi_Y: Z \to Y$ be surjective maps from the finite set $Z$ into $X$ and $Y$, respectively, such that $|\varphi_X^* F_X(\tau) - \varphi_Y^* F_Y(\tau)| < \varepsilon$ for all $\tau \in \text{pow}(Z)$. Then, by Theorem 4.1,
\[
d_D(D_k(\varphi_X^* F_X), D_k(\varphi_Y^* F_Y)) < \varepsilon.
\]
for all $k \in \mathbb{N}$. Now apply Corollary 4.1 and conclude by letting $\varepsilon$ approach $d_F(X, Y)$. \hfill \Box

5. FILTRATIONS ARISING FROM METRIC SPACES: RIPS AND ČECH

Recall [BB10] that for two compact metric spaces $(X, d_X)$ and $(Y, d_Y)$, a correspondence between them is any subset $R$ of $X \times Y$ such that the natural projections $\pi_X: X \times Y \to X$ and $\pi_Y: X \times Y \to Y$ are such that $\pi_X(R) = X$ and $\pi_Y(R) = Y$. The distortion of any such correspondence is given by
\[
dis(R) := \sup_{(x,y), (x',y') \in R} |d_X(x, x') - d_Y(y, y')|.
\]
Then, Gromov-Hausdorff distance between $(X, d_X)$ and $(Y, d_Y)$ is defined as
\[
d_{GH}(X, Y) := \frac{1}{2} \inf_R \text{dis}(R),
\]
where the infimum is taken over all correspondences $R$ between $X$ and $Y$.

5.1. **The Rips filtration.** Recall the definition of the *Rips filtration* of a finite metric space $(X, d_X)$: for $\sigma \in \text{pow}(X)$,
\[
F_X^R(\sigma) = \text{diam}_X(\sigma) := \max_{x, x' \in X} d_X(x, x').
\]

The following theorem was first proved in [CCSG*09]. A different proof (also applicable to compact metric spaces) relying on the interleaving distance and multivalued maps was given in [CDSO14]. Yet another different proof avoiding multivalued maps is given in [CM16].

**Theorem 5.1.** For all finite metric spaces $X$ and $Y$, and all $k \in \mathbb{N}$,
\[
d_D(D_k(F_X^R), D_k(F_Y^R)) \leq 2 d_{GH}(X, Y).
\]

A different proof of Theorem 5.1 can be obtained by combining Theorem 4.2 and Proposition 5.1 below.
Proof of Proposition 5.1. For all finite metric spaces $X$ and $Y$,
\[ d_F(F_X^R, F_Y^R) \leq 2d_{\mathcal{GH}}(X, Y). \]

Proof of Proposition 5.1. Let $X$ and $Y$ be s.t. $d_{\mathcal{GH}}(X, Y) < \eta$, and let $R \subseteq X \times Y$ be a surjective relation with $|d_X(x, x') - d_Y(y, y')| \leq 2\eta$ for all $(x, y), (x', y') \in R$. Consider the parametrization $Z = R$, and $\varphi_X = \pi_1 : Z \to X$ and $\varphi_Y = \pi_2 : Z \to Y$, then
\[ |d_X(\varphi_X(t), \varphi_X(t')) - d_Y(\varphi_Y(t), \varphi_Y(t'))| \leq 2\eta \]
for all $t, t' \in Z$. Pick any $\tau \in Z$ and notice that
\[ \varphi^*_X F^R_X(\tau) = F^R_X(\varphi_X(\tau)) = \max_{t, t' \in \tau} d_X(\varphi_X(t), \varphi_X(t')). \]

Now, similarly, write
\[ \varphi^*_Y F^R_Y(\tau) = \max_{t, t' \in \tau} d_Y(\varphi_Y(t), \varphi_Y(t')) \leq \max_{t, t' \in \tau} d_X(\varphi_X(t), \varphi_X(t')) + 2\eta = \varphi^*_X F^R_X(\tau) + 2\eta, \]
where the last inequality follows from (2). The proof follows by interchanging the roles of $X$ and $Y$. \hfill \Box

5.2. The Čech filtration. Another interesting and frequently used filtration is the Čech filtration: for each $\sigma \in \text{pow}(X)$,
\[ F^C_X(\sigma) := \text{rad}_X(\sigma) = \min_{p \in X} \max_{x \in \sigma} d_X(x, p). \]
That is, the filtration value of each simplex corresponds to its circumradius.

Proposition 5.2. For all finite metric spaces $X$ and $Y$,
\[ d_F(F^W_X, F^W_Y) \leq 2d_{\mathcal{GH}}(X, Y). \]

Again, as a corollary of Theorem 4.2 and Proposition 5.2 we have the following

Theorem 5.2. For all finite metric spaces $X$ and $Y$, and all $k \in \mathbb{N}$,
\[ d_D(D_k(F^C_X), D_k(F^C_Y)) \leq 2d_{\mathcal{GH}}(X, Y). \]

A proof of this theorem via the interleaving distance and multi-valued maps has appeared in [CDSO14]. Another proof avoiding multivalued maps is given in [CM16].

Proof of Proposition 5.2. The proof is similar to that of Proposition 5.1. Pick any $\tau \in Z$, then
\[ \varphi^*_X F^W_X(\tau) = F^W_X(\varphi_X(\tau)) = \min_{p \in X} \max_{t \in \tau} d_X(p, \varphi_X(t)) = \max_{t \in \tau} d_X(p_\tau, \varphi_X(t)) \]
for some $p_\tau \in X$. Let $t_\tau \in Z$ be s.t. $\varphi_X(t_\tau) = p_\tau$, and from the above obtain
\[ \varphi^*_X F^W_X(\tau) = \max_{t \in \tau} d_X(\varphi_X(t_\tau), \varphi_X(t)). \]

The version in [CDSO14] applies to compact metric spaces.
Now, similarly, write

\[ \varphi^*_Y F^W_Y(\tau) = \min_{q \in Y} \max_{t \in \tau} d_Y(q, \varphi_Y(t)) \leq \max_{t \in \tau} d_Y(\varphi_Y(t), \varphi_Y(t)) \leq \max_{t \in \tau} d_X(\varphi_X(t), \varphi_X(t)) + 2\eta = \varphi^*_X F^W_X(\tau) + 2\eta, \]

where the last inequality follows from (2). The proof follows by interchanging the roles of \(X\) and \(Y\). \qed

6. Discussion

It seems possible to extend some of these ideas to the case of not necessarily finite filtered spaces.

References

[Ban12] BIRS station. Banff. \url{http://www.birs.ca/events/2012/5-day-workshops/12w5081/videos/watch/2012101612012}.

[BBI01] D. Burago, Y. Burago, and S. Ivanov. A Course in Metric Geometry, volume 33 of AMS Graduate Studies in Math. American Mathematical Society, 2001.

[Car14] Gunnar Carlsson. Topological pattern recognition for point cloud data. Acta Numerica, 23:289368, 2014.

[CCSG+09] F. Chazal, D. Cohen-Steiner, L. Guibas, F. Mémoli, and S. Oudot. Gromov-Hausdorff stable signatures for shapes using persistence. In Proc. of SGP, 2009.

[CDS10] Gunnar Carlsson and Vin De Silva. Zigzag persistence. Foundations of computational mathematics, 10(4):367–405, 2010.

[CDSO14] Frédéric Chazal, Vin De Silva, and Steve Oudot. Persistence stability for geometric complexes. Geometriae Dedicata, 173(1):193–214, 2014.

[CM16] Samir Chowdhury and Facundo Mémoli. Persistent homology of asymmetric networks: An approach based on dowker filtrations. \textit{arXiv preprint arXiv:1608.05432}, 2016.

[CSEM06] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In SCG ’06: Proceedings of the twenty-second annual symposium on Computational geometry, pages 119–126, New York, NY, USA, 2006. ACM.

[EH10] Herbert Edelsbrunner and John Harer. Computational Topology - an Introduction. American Mathematical Society, 2010.

[Qui73] Daniel Quillen. Higher algebraic k-theory: I. In Higher K-theories, pages 85–147. Springer, 1973.

E-mail address: memoli@math.osu.edu

(F. Mémoli) Department of Mathematics and Department of Computer Science and Engineering, The Ohio State University. Phone: (614) 292-4975, Fax: (614) 292-1479.