Posets, metric spaces, and topological data analysis

Rick Jardine

University of Western Ontario

September 16, 2020
Basic setup

$X$ data set: $Z$ metric space, $X \in D(Z) = \text{finite subsets of } Z$.

• $P_s(X)$ = poset of subsets $\sigma \subset X$ such that $d(x, y) \leq s$ for all $x, y \in \sigma$. $s \in [0, \infty)$.

$P_s(X)$ is the poset of simplices of **Vietoris-Rips complex** $V_s(X)$.

The **nerve** $BP_s(X)$ is **barycentric subdivision** of $V_s(X)$ (same homotopy type).

We have poset inclusions

$$\sigma : P_s(X) \subset P_t(X), \ s \leq t,$$

$P_0(X) = X$, and $P_t(X) = \mathcal{P}(X)$ (all subsets of $X$) for $t$ suff large.

• $k \geq 0$: $P_{s,k}(X) \subset P_s(X)$ subposet of simplices $\sigma$ such that each element $x \in \sigma$ has at least $k$ distinct neighbours $y$ such that $d(x, y) \leq s$.

$P_{s,k}(X)$ is the poset of simplices of the **degree Rips complex** $L_{s,k}(X)$. $BP_{s,k}(X)$ is the **barycentric subdivision** of $L_{s,k}(X)$.
The nerve $BC$ of a category $C$ is a simplicial set with $n$-simplices $BC_n$ given by the set of strings of arrows

$$a_0 \to a_1 \to \cdots \to a_n$$

of length $n$ in $C$, equivalently functors $n \to C$, where

$$n = \{0, 1, \ldots, n\},$$

with the obvious poset structure.

Composition with the functors $\theta : m \to n$ defines the simplicial structure of $BC$.

**Examples:**

1) $Bn = \Delta^n$, the standard $n$-simplex in simplicial sets.

2) $BG = K(G, 1)$ for a group $G$, classifies principal $G$-bundles.

$BC$ is also called the **classifying space** of $C$. 

Rick Jardine  
Posets, metric spaces, and topological data analysis
Theorem 1 (Rips stability).

Suppose \( X \subset Y \) in \( D(Z) \) such that \( d_H(X, Y) < r \). There is a homotopy commutative diagram (homotopy interleaving)

\[
P_s(X) \xrightarrow{\sigma} P_{s+2r}(X) \\
P_s(Y) \xrightarrow{\sigma} P_{s+2r}(Y)
\]

Corollary 2 (Stability for persistence invariants).

Same assumptions as Theorem 1. There are commutative diagrams

\[
H_k(V_s(X)) \xrightarrow{\sigma} H_k(V_{s+2r}(X)) \\
H_k(V_s(Y)) \xrightarrow{\sigma} H_k(V_{s+2r}(Y))
\]

There is a corresponding statement for \( \pi_0 \) (clusters).
Sketch proof

$y \in Y$: there is $\theta(y) \in X$ st. $d(y, \theta(y)) < r$ (from $d_H(X, Y) < r$).

$x \in X$: $\theta(x) = x$.

$\sigma = \{y_1, \ldots, y_k\}$ in $P_s(Y)$, then

$$\sigma \cup \theta(\sigma) = \{y_1, \ldots, y_k, \theta(y_1), \ldots, \theta(y_k)\} \in P_{s+2r}(Y)$$

and there are homotopies (natural transformations)

$$\sigma \subseteq \sigma \cup \theta(\sigma) \supseteq \theta(\sigma).$$

between poset morphisms $P_s(Y) \to P_{s+2r}(Y)$. 

Rick Jardine

Posets, metric spaces, and topological data analysis
A natural transformation $h$ between functors $f, g : C \to D$ is a diagram of functors

\[
\begin{array}{ccc}
C & \xrightarrow{f} & D \\
\downarrow{i_0} & & \downarrow{g} \\
C \times 1 & \xrightarrow{h} & D \\
\uparrow{i_1} & & \\
C & \xrightarrow{f} & D
\end{array}
\]

where $1 = \{0 \leq 1\}$, $i_\epsilon(a) = (a, \epsilon)$.

\[
B(C \times 1) \cong BC \times B1 = BC \times \Delta^1
\]
Stability: degree Rips

\[ D(Z) = \text{finite subsets of a metric space } Z. \]

**Theorem 3.**

Suppose \( X \subset Y \) in \( D(Z) \) such that \( d_H(X_{dis}^{k+1}, Y_{dis}^{k+1}) < r \). There is a homotopy commutative diagram

\[
\begin{array}{ccc}
P_{s,k}(X) & \xrightarrow{\sigma} & P_{s+2r,k}(X) \\
\downarrow i & & \downarrow i \\
P_{s,k}(Y) & \xrightarrow{\sigma} & P_{s+2r,k}(Y)
\end{array}
\]

\( X_{dis}^{k+1} \) is the set of subsets \( \sigma \subset X \) having \( k + 1 \) elements.

\( X_{dis}^{k+1} \subset Z^{k+1} \).
Equivalence up to shift

**NB:** $V_*(X) := BP_*(X)$ for a while.

Suppose that $X \subset Y$ in $D(Z)$ and we have a homotopy interleaving

$$
\begin{array}{c}
V_s(X) \xrightarrow{\sigma} V_{s+2r}(X) \\
\downarrow i \quad \theta \quad \downarrow i \\
V_s(Y) \xrightarrow{\sigma} V_{s+2r}(Y)
\end{array}
$$

1) $i : \pi_0 V_*(X) \rightarrow \pi_0 V_*(Y)$ is a 2r-monomorphism: if $i([x]) = i([y])$ in $\pi_0 V_s(Y)$ then $\sigma([x]) = \sigma([y])$ in $\pi_0 V_{s+2r}(X)$

2) $i : \pi_0 V_*(X) \rightarrow \pi_0 V_*(Y)$ is a 2r-epimorphism: given $[y] \in \pi_0 V_s(Y)$, $\sigma([y]) = i([x])$ for some $[x] \in \pi_0 V_{s+2r}(X)$.

3) All $i : \pi_n(V_*(X), x) \rightarrow \pi_n(V_*(Y), i(x))$ are 2r-isomorphisms.

The map $i : V_*(X) \rightarrow V_*(Y)$ is a 2r-equivalence of systems.
A **system** of spaces is a functor $X : [0, \infty) \to \text{sSet}$, aka. a **diagram** of simplicial sets with index category $[0, \infty)$.

A **map** of systems $X \to Y$ is a natural transformation of functors defined on $[0, \infty)$. 

$s\text{Set}^{[0,\infty)}$ is the category of systems and natural transformations.

**Examples**

1) The functor $V_*(X)$, $s \mapsto V_s(X) = BP_s(X)$ is a system of spaces, for a data set $X \subset Z$.

2) If $X \subset Y$ in $D(Z)$, the induced maps $P_s(X) \to P_s(Y)$, $V_s(X) \to V_s(Y)$ define maps of systems $P_\ast(X) \to P_\ast(Y)$ (posets) and $V_\ast(X) \to V_\ast(Y)$ (spaces).
Homotopy theories of systems: oldest is the **projective structure** (Bousfield-Kan, 1972):

A map \( f : X \to Y \) is a **weak equivalence** (resp. **fibration**) if each map \( X_s \to Y_s \) is a weak equiv. (resp. Kan fibration) of simplicial sets.

A map \( A \to B \) is a **projective cofibration** if it has the left lifting property with respect all maps which are trivial fibrations.

**Lemma 4.**

*Suppose that* \( X \subset Y \) *in* \( D(Z) \). *Then* \( V_*(X) \to V_*(Y) \) *is a projective cofibration.*

The map \( V_*(X) \to V_*(Y) \) is also a **sectionwise** cofibration, i.e. all maps \( V_s(X) \to V_s(Y) \) are monomorphisms.

Fibrations and weak equivalences for the projective structure are defined sectionwise.
Controlled equivalences

Suppose that \( f : X \to Y \) is a map of systems. Say that \( f \) is an \( r \)-equivalence if

1) the map \( f : \pi_0(X) \to \pi_0(Y) \) is an \( r \)-isomorphism of systems of sets

2) the maps \( f : \pi_k(X_s, x) \to \pi_k(Y_s, f(x)) \) are \( r \)-isomorphisms of systems of groups, for all \( s \geq 0, x \in X_s, k \geq 1 \).

Observation: Suppose given a diagram of systems

\[
\begin{array}{ccc}
X_1 & \xrightarrow{f_1} & Y_1 \\
\searrow & \mathring{\cong} & \searrow \\
X_2 & \xrightarrow{f_2} & Y_2
\end{array}
\]

Then \( f_1 \) is an \( r \)-equivalence iff \( f_2 \) is an \( r \)-equivalence.

Examples: Stability results. A sectionwise equivalence is a 0-equivalence.

A controlled equivalence is a map which is an \( r \)-equivalence for some \( r \geq 0 \).
Lemma 5.

Suppose given a commutative triangle

\[
\begin{array}{ccc}
X & \xrightarrow{f} & Y \\
\downarrow{h} & & \downarrow{g} \\
Z & & 
\end{array}
\]

If one of the maps is an \(r\)-equivalence, a second is an \(s\)-equivalence, then the third map is a \((r + s)\)-equivalence.

Proof.

Set theory.
Theorem 6.

Suppose that \( i : A \to B \) is a sectionwise cofibration and an \( r \)-equivalence, and suppose given a pushout

\[
\begin{array}{ccc}
A & \to & C \\
\downarrow & & \downarrow i_* \\
B & \to & D
\end{array}
\]

Then \( i_* \) is a sectionwise cofibration and a \( 2r \)-equivalence.

Sketch (Whitehead theorem): There is a \( 2r \)-interleaving

\[
\begin{array}{ccc}
A_s & \to & FA_{s+2r} \\
\downarrow & \nearrow & \downarrow \\
B_s & \to & FB_{s+2r}
\end{array}
\]

for a sectionwise fibrant model of \( i \). The class of cofibrations admitting \( 2r \)-interleavings is closed under pushout.
(A'): Suppose given a commutative diagram

\[
\begin{array}{ccc}
A & \rightarrow & C \\
\downarrow & & \downarrow \\
B & \rightarrow & \ \\
\end{array}
\]

If one of the maps is an \(r\)-equivalence, another is an \(s\)-equivalence, then the third is an \((r + s)\)-equivalence.

(B): The composite of two cofibrations is a cofibration. Any isomorphism is a cofibration.

(C'): Cofibrations are closed under pushout. Given a pushout

\[
\begin{array}{ccc}
A & \rightarrow & C \\
\downarrow & & \downarrow \\
i & & i_* \\
B & \rightarrow & D \\
\end{array}
\]

with \(i\) a cofibration and \(r\)-equivalence, then \(i_*\) is a cofibration and a \(2r\)-equivalence.
(D): For any object $A$ there is at least one cylinder object $A \otimes \Delta^1$.

(E): All objects are cofibrant.

This is an adjusted list of axioms for a category of cofibrations structure — works for projective or sectionwise cofibrations.

Lemma 7 (left properness).

Suppose given a pushout

$$
\begin{array}{ccc}
A & \xrightarrow{u} & C \\
\downarrow{\scriptstyle i} & & \downarrow{} \\
B & \xrightarrow{u_*} & D
\end{array}
$$

where $i$ is a cofibration and $u$ is an r-equivalence. Then $u_*$ is a 2r-equivalence.

There is also a patching lemma.
An extended pseudo-metric space (ep-metric space) \((X, D)\) is a set \(X\) and a function \(D : X \times X \to [0, \infty]\) such that

1) \(D(x, x) = 0\),
2) \(D(x, y) = D(y, x)\),
3) \(D(x, z) \leq D(x, y) + D(y, z)\).

- Can have distinct \(x, y\) such that \(D(x, y) = 0\) ("pseudo").
- Can have \(u, v\) such that \(D(u, v) = \infty\) ("extended").

Every metric space \((Y, d)\) is an ep-metric space via composition

\[ Y \times Y \xrightarrow{d} [0, \infty) \subset [0, \infty]. \]

A morphism \(f : (X, d_X) \to (Y, d_Y)\) of ep-metric spaces is a function \(f : X \to Y\) such that

\[ d_Y(f(x), f(y)) \leq d_X(x, y) \text{ (non-expanding)}. \]

\(ep – Met\) is the category of ep-metric spaces and their morphisms.
Quotient construction

$(X, d)$ an ep-metric space and $p : X \to Y$ a surjective function.

For $x, y \in Y$ set

$$D(x, y) = \inf_P \sum_{i=0}^k d(x_i, y_i),$$

“Polygonal path” $P :$ pairs $(x_i, y_i)$ in $X$ with $x = p(x_0)$, $y = p(x_k)$, $p(y_i) = p(x_{i+1})$.

For $x, y \in X$, $P : x, y$ is path from $p(x)$ to $p(y)$, so $D(p(x), p(y)) \leq d(x, y)$.

Polygonal paths concatenate, so $D(x, z) \leq D(x, y) + D(y, z)$.

$D(x, x) = 0$ and $D(x, y) = D(y, x)$.

Quotient map $p : (X, d) \to (Y, D)$ satisfies universal property.
1) Suppose \((X_i, d_i), \ i \in I\) is a set of ep-metric spaces. There is an ep-metric \(D\) on \(\bigsqcup_i X_i\), with
\[
D(x, y) = \begin{cases} 
  d_i(x, y) & \text{if } x, y \in X_i, \\
  \infty & \text{if } x, y \text{ are in different summands}.
\end{cases}
\]

\(\bigsqcup_i (X_i, d_i)\) is a \textbf{coproduct} in \(ep – Met\).

2) Suppose given morphisms \(f, g : (X, d_X) \to (Y, d_Y)\) in \(ep – Met\). Form the set theoretic coequalizer
\[
X \xrightarrow{f} Y \xrightarrow{g} \xrightarrow{p} C,
\]

Then \(p\) is a surjective function, and we give \(C\) the quotient ep-metric \(D\).

\[(X, d_X) \xrightarrow{(f, g)} (Y, d_Y) \xrightarrow{p} (C, D)\]

is a \textbf{coequalizer} in \(ep – Met\).
Suppose given subspaces $X, Y \subset Z$ of a finite metric space $(Z, d)$ such that $Z = X \cup Y$.

$X \cup_{em} Y$ is **pushout** in $ep - Met$.

The map $X \cup_{ep} Y \to Z$ is the identity on the underlying sets.

The metric $D$ on $X \cup_{ep} Y$ is defined by

$$D(x, y) = \inf_P \left\{ \sum_i d(x_i, x_{i+1}) \right\},$$

indexed over sequences $P : x = x_0, x_1, \ldots, x_k = y$ such that each pair $(x_i, x_{i+1})$ is either in $X$ or in $Y$.

**Fact**: The induced map

$$\pi_0(V_s(X, d) \cup V_s(Y, d)) \to \pi_0 V_s(X \cup_{ep} Y, D)$$

is a bijection.

**Question**: What about homotopy/homology groups?
$sSet^{[0,\infty]}$ is the diagram category of functors $[0, \infty] \to sSet$.

Suppose that $K$ is a simplicial set and $s \in [0, \infty]$.

$$L_s K_t = \begin{cases} K & \text{if } t \geq s, \\ \emptyset & \text{if } t < s. \end{cases}$$

**Fact:** $\text{hom}(L_s K, X) = \text{hom}(K, X_s)$ for $X : [0, \infty] \to sSet$

A **simplex** of $X$ is a morphism $\sigma : L_s \Delta^n \to X$ (an $n$-simplex of $X_s$). A **morphism of simplices** is a commutative diagram

$$
\begin{array}{ccc}
L_t \Delta^m & \xrightarrow{\theta} & L_s \Delta^n \\
\downarrow{\tau} & & \downarrow{\sigma} \\
X & & X_s \\
\end{array}
\quad
\begin{array}{ccc}
\Delta^m & \xrightarrow{\theta} & \Delta^n \\
\downarrow{\tau} & & \downarrow{\sigma} \\
X_t & \leftarrow & X_s \\
\end{array}
$$

The **simplex category** is denoted by $\Delta/X$. 

Rick Jardine  
Posets, metric spaces, and topological data analysis
Write $U_s^n$ for the metric space structure on $n = \{0, 1, \ldots, n\}$ with $d(i, j) = s$.

If $\theta : m \rightarrow n$ is a poset map and $s \leq t$, then $\theta$ induces an ep-metric space map $\theta : U_t^m \rightarrow U_s^n$, since

$$d(\theta(i), \theta(j)) \leq s \leq t = d(i, j), \text{ if } i \neq j.$$ 

The assignment which takes a simplex $\sigma : L_s \Delta^n \rightarrow X$ to the space $U_s^n$ defines a functor $\Delta/X \rightarrow ep – Met$. Then

$$\text{Re}(X) := \lim_{\rightarrow \text{L}_s \Delta^n \rightarrow X} U_s^n$$

is the **realization** of $X$ in $ep – Met$.

If $Y \in ep – Met$ then $S(Y)$ is the diagram with

$$S(Y)_{s,n} = \text{hom}(U_s^n, Y).$$

**Singular functor** $S : ep – Met \rightarrow sSet^{[0,\infty]}$ is **right adjoint** to $\text{Re}$. 
Suppose that $X : [0, \infty] \to \text{sSet}$ is a diagram. Then

$$\lim_{s \to \infty} X_s = X_\infty$$

because $\infty$ is terminal in $[0, \infty]$.

**Lemma 8.**

*Given $X : [0, \infty] \to \text{sSet}$, then $\text{Re}(X) \cong (X_\infty, D)$, where

$$D(x, y) = \inf_P \left\{ \sum_{i=0}^{k} s_i \right\}$$

where $P$ is a polygonal path consisting of 1-simplices $x_i \to y_i$ in $X_{s_i}$, with $x_0 \leftrightarrow x$, $x_k \leftrightarrow y$, and $y_i, x_{i+1}$ have the same image in $X_\infty$.***
Examples

$(X, d)$ a totally ordered finite ep-metric space.

The Vietoris-Rips complex $V_s(X)$ is a simplicial set with $n$-simplices

$$x_0 \leq x_1 \leq \cdots \leq x_n$$

with $d(x_i, x_j) \leq s$ for all $i, j$. $BP_s(X)$ is the barycentric subdivision of $V_s(X)$. Here, $s \leq \infty$.

Lemma 9.

$X$ a totally ordered finite ep-metric space. Then $\text{Re}(V_*(X)) \cong X$.

Proof.

$D(x, y) = \inf_P \sum d(x_i, x_{i+1})$, indexed over sequences $x = x_0, x_1, \ldots, x_k = y$. Then $d(x, y) \leq D(x, y)$ by the triangle identity.

$D(x, y) \leq d(x, y)$ because $(x, y)$ is a path in some $X_s$. 

Rick Jardine

Posets, metric spaces, and topological data analysis
(X, d) a totally ordered finite ep-metric space. The canonical map η : V(X) → S(Re(V(X))) has the form

η : V(X) → S(X),

where η takes the simplex x₀ ≤ x₁ ≤ ⋅⋅⋅ ≤ xₙ to the sequence (x₀, x₁, ⋅⋅⋅, xₖ) (forgets the ordering — the sequence is a “bag of words”).

Theorem 10.

(X, d) a totally ordered finite ep-metric space. Then each map

η : Vₛ(X) → S(X)ₛ

is a weak equivalence of simplicial sets.

Proof uses simplicial approximation techniques. Show that

BNVₛ(X) → BNS(X)ₛ and π : sd S(X)ₛ → BNS(X)ₛ are weak equivalences.
(X, \ d) is a finite ep-metric space.
Choose a \textbf{neighbourhood set} \( N_x \) for each \( x \in X \).
Set \( r_x = \max_{y \in N_x} d(x, y) \).
e.g. \( N_x \) is set of \( k \) nearest neighbours if \( X \) is totally ordered, for some \( k \).
Set \( (U_x, d_x) = \bigvee_{y \in N_x} (\{x, y\}, d) \) in \( \text{ep} - \text{Met} \).
Then \( d_x(y_1, y_2) = d(x, y_1) + d(x, y_2) \) for \( y_1, y_2 \in N_x \).
Extend to an ep-metric \( d_x \) on \( X \) by setting \( d_x(y, z) = \infty \) if either \( y \) or \( z \) is outside of \( U_x \).
We have inclusions \( X \subset V(X, d_x), x \in X \). Form iterated pushout
\[ V(X, N) = \bigvee_X V(X, d_x) \simeq \bigvee_X S(X, d_x). \]
The metrics \( d_x \) can be rescaled, but the diagram \( V(X, N) \) is “the” \textbf{UMAP complex}.
\textbf{UMAP algorithm:} apply TDA machinery (e.g. \( \pi_0 \)) to \( V(X, N) \).
$(X, d)$ is a finite totally ordered ep-metric space, with
neighbourhoods $N = \{ N_x, \ x \in X \}$.

$$\phi : V(X, N) = \bigvee_X V(X, d_x) \rightarrow V(X)$$
canonical map.

$(x, y)$ in $X$ is a **neighbourhood pair** if $y \in N_x$ or $x \in N_y$.

Graph $\Gamma(X, N) \overset{i}{\subset} V(X)$ with vertices $X$ and edges all nbhd pairs.

**Lemma 11.**

$$\pi_0 V(X, N)_s \overset{\sim}{\rightarrow} \pi_0 \Gamma(X, N)_s$$

If all 1-simplices of $V_s(X)$ are nbhd pairs, then $i_*$ is an iso.

**Example:** $N_x = k$-nearest neighbours, $r_x = \max_{y \in N_x} d(x, y)$,
$s < r_x$ for all $x$. 
Fact: $V(X, N)_{\infty}$ is a big wedge of circles.

$V(X, d_{X}) = \Delta^{X} = \Delta^{N}$ for $N + 1 = |X|$, so $V(X, N)_{\infty} = \vee_{N} \Delta^{n} (N + 1 \text{ summands})$.

Define $N \to \Delta^{N} = X_{i}, 0 \leq i \leq k$, $Y = \vee_{N} X_{i}$ (iterated pushout).
Each $X_{i}$ is contractible, so $Y / X_{0} \simeq Y$, and

$$Y / X_{0} = (X_{1}/N) \vee \cdots \vee (X_{k}/N) = (\Delta^{N} / N) \vee \cdots \vee (\Delta^{N} / N)$$

and each

$$\Delta^{N} / N \simeq \Sigma N \simeq \Sigma (S^{0} \vee \cdots \vee S^{0}) \ (N \text{ summands, } N \text{ pointed by } 0)$$

$$\simeq S^{1} \vee \cdots \vee S^{1}.$$

Consequence: $V(X, N)_{\infty} \simeq \bigvee_{N-1}^{N} S^{1} \ (N = |X| - 1)$. 
References

J.F. Jardine.
Metric spaces and homotopy types.

J.F. Jardine.
Persistent homotopy theory.

Leland McInnes and John Healy.
Accelerated hierarchical density based clustering.

D.I. Spivak.
Metric realization of fuzzy simplicial sets.