Fuzzy sets, presheaves, and topological data analysis

Rick Jardine

University of Western Ontario

November 12, 2020

Fuzzy sets

Classical definition (ca. 1965): **Fuzzy sets** are functions

$$\phi : X \to [0,1].$$

Given $\psi : Y \to [0,1]$, a **morphism**

$$f : \phi \to \psi$$

of fuzzy sets consists of a function $f : X \to Y$ and a relation (homotopy) $\phi \leq \psi \cdot f$ of functions taking values in $[0,1]$.

i.e. $\phi(x) \leq \psi(f(x))$ for all $x \in X$.

Fuzz is the category of fuzzy sets.

Locales

$[0, 1]$ is a locale ...

A **locale** (also **frame**) L is a poset with infinite joins (unions) and finite meets (intersections), in which finite meets distribute over all joins. Isbell (1972).

NB: L has a terminal object (empty meet), an initial object (empty join), and infinite meets.

A **morphism** of locales $L_1 \to L_2$ is a poset morphism which preserves meets and joins (hence preserves initial and terminal objects).

Examples

1) Any closed interval $[a,b] \subset \mathbb{R}$ (standard ordering) is a locale.
$[0,\infty]$ is a locale.

The scaling isomorphism $[0,1] \to [a,b]$, defined by

$$t \mapsto t \cdot b + (1-t) \cdot a$$

is an isomorphism of locales.

2) $op|_X =$ open subsets of a topological space X is a locale.

3) The opposite poset $[a,b]^{op}$ is a locale.

4) $[0,\infty]^{op}$ is a locale.

5) Power set $\mathcal{P}(X)$ on a set X.

Barr's definition (1986)

Michael Barr [1]: A **fuzzy set**, over a locale L, is a function

$$\phi : X \to L.$$

$\psi : Y \to L$ is another function: a **morphism** $f : \phi \to \psi$ of the corresponding fuzzy sets consists of a function $f : X \to Y$ such that

$$\phi(x) \leq \psi(f(x))$$

for all $x \in X$.

Fuzz(L) is the category of fuzzy sets over L.

Sheaves

Every locale L has a Grothendieck topology, for which the covering families of a are sets of elements $b_i \leq a$ such that $\vee_i \; b_i = a$.

There are associated categories **Pre**(L) and **Shv**(L) of presheaves and sheaves, respectively.

A (set valued) **presheaf** is a contravariant functor $L^{op} \to \mathbf{Set}$.

A **sheaf** is a presheaf $F : L^{op} \to \mathbf{Set}$ such that the diagram

$$F(a) \to \prod_i \; F(b_i) \rightrightarrows \prod_{i,j} \; F(b_i \wedge b_j)$$

is an equalizer for all cov. families $b_i \leq a$, $a \in L$ (patching condition).

Morphisms of presheaves and/or sheaves are just natural transformations. **Shv**(L) is a full subcategory of **Pre**(L).

Fuzzy sets beget sheaves

Suppose $\phi : X \to L$ is a fuzzy set over a locale L, $a \in L$, and write

$$L_{\geq a} = \{x \in L \mid x \geq a\}.$$

Form the pullback

$$\begin{array}{ccc} \phi^{-1}(L_{\geq a}) & \to & X \\ \downarrow & & \downarrow \phi \\ L_{\geq a} & \longrightarrow & L \end{array}$$

The assignment $T(\phi)(0_+) = *$, and

$$a \mapsto T(\phi)(a) := \phi^{-1}(L_{\geq a}),\ a \in L,$$

defines a sheaf $T(\phi)$ on $L_+ = L \sqcup \{0_+\}$ (new initial elt. 0_+).

In effect, if $a_i \leq b$ covers b then $L_{\geq b} = \cap_i L_{\geq a_i}$

Sheaves of monomorphisms

Given $\phi : X \to L$, the restrictions $\phi^{-1}(L_{\geq b}) \to \phi^{-1}(L_{\geq a})$ are monomorphisms for all $a \leq b$ in L, so that $T(\phi)$ is a **sheaf of monomorphisms** on L_+.

$\mathbf{Mon}(L_+)$ is the category of sheaves of monomorphisms on L_+.

We have defined a functor

$$T : \mathbf{Fuzz}(L) \to \mathbf{Mon}(L_+).$$

Theorem 1 (Barr, 1986).

There is an equivalence of categories

$$T : \mathbf{Fuzz}(L) \leftrightarrows \mathbf{Mon}(L_+) : S$$

Write $i \in L \subset L_+$ for the original initial object of L.

Def: F a sheaf of monics on L_+: $F(i)$ is the **generic fibre** of F.

All restriction maps $F(a) \to F(i)$ are monomorphisms for $a \in L$.

For each $x \in L(i)$, there is a maximum (lub) b such that x is in the image of $F(b) \to F(i)$.

For $F \in \mathbf{Mon}(L_+)$,

$$S(F) : F(i) \to L$$

is the function which sends x to b.

X finite metric space, with a listing $X \cong \{0, 1, \dots, N\} = \mathbf{N}$ (finite ordinal number).

Choose R such that $d(x, y) < R$ for all $x, y \in X$.

Setting $\phi(\sigma) = \max_{i,j}\ d(x_i, x_j)$ for a simplex $\sigma = \{x_0, x_1, \dots, x_k\}$ defines a fuzzy set

$$\phi : \Delta^X_k := \Delta^N_k \to [0, R]^{op}.$$

The assoc. sheaf of monomorphisms $T(\phi)$ on $[0, R]^{op}_+$ has

$$T(\phi)(s) = \phi^{-1}([0, R]^{op}_{\geq s}) = \phi^{-1}([0, s]) = V_s(X)_k,$$

set of k-simplices of the **Vietoris-Rips complex** $V_s(X)$.

The simplicial sheaf of monomorphisms associated to the simplicial fuzzy set $\Delta^X \to [0, R]^{op}$ is the system of Vietoris-Rips complexes $s \mapsto V_s(X)$, $0 \leq s \leq R$.

Also $V_{0_+}(X) = *$ for the initial object $0_+ \in [0, R]^{op}_+$.

Presheaves of monomorphisms

A presheaf $F : L_+^{op} \to \mathbf{Set}$ such that

1) $F(0_+) = *$, and
2) all $a \leq b$ in L induce monomorphisms $F(b) \to F(a)$

is a **presheaf of monomorphisms**.

$\mathbf{Mon}_p(L_+)$ is the category of **presheaves of monomorphisms**.

Associated sheaf

Lemma 2.

Suppose that $L = [a, b]$. The covering sieves for $s \in L$ are the families of all r such that $r < s$ or such that $r \leq s$.

Consequence: A presheaf F on L_+ is a sheaf if and only if $F(0) = *$ and the map

$$\eta : F(s) \to \varprojlim_{0<r<s} F(r) =: LF(s) \tag{1}$$

is an isomorphism for all $a \in L$ with a not initial.

LF is the *separated presheaf* associated to F.

Lemma 3.

Suppose that $L = [a, b]$, and $F \in \mathbf{Mon}_p(L_+)$. Then F is separated, so LF is a sheaf and $\eta : F \to LF$ is the associated sheaf map. LF is a sheaf of monomorphisms (fuzzy set).

Colimits

Suppose E is a presheaf on L_+. The epi-monic factorizations of the maps $E(s) \to E(i)$ for $s \in L$ define $\mathrm{Im}(E)(s) \subset E(i)$, with

$$\begin{array}{ccccc} E(t) & \longrightarrow & \mathrm{Im}(E)(t) & & \\ \downarrow & & \downarrow & \searrow & \\ & & & & E(i) \\ & & & \nearrow & \\ E(s) & \longrightarrow & \mathrm{Im}(E)(s) & & \end{array}$$

for $s \leq t$. Set $\mathrm{Im}(E)(0) = *$.

If $E \in \mathbf{Mon}_p(L_+)$, then all $E(s) \to \mathrm{Im}(E)(s)$ are isomorphisms.

$\mathrm{Im}(E)$ is a presheaf of monomorphisms, and there is a natural bijection

$$\hom_{\mathbf{Mon}_p(L_+)}(\mathrm{Im}(E), F) \cong \hom(E, F),$$

Corollary 4.

$\mathbf{Mon}_p(L_+)$ *and* $\mathbf{Mon}(L_+) \simeq \mathbf{Fuzz}(L)$ *are co-complete.*

Stalks

For a sheaf F on $L = [a, b]$ and $t \in L - \{a\}$, the stalk F_t is def. by

$$F_t = \varinjlim_{t<s} F(s).$$

Example: X finite metric space. $s \mapsto V_s(X)$ Vietoris-Rips simplicial sheaf on $[0, R]_+^{op}$ ($d(x, y) \leq R$ for all $x, y \in X$).

The stalk $V(X)_s$ for $s \in (0, R]^{op}$ is defined by

$$V(X)_s = \varinjlim_{t<s} V_t(X), \text{ for std. order in } [0, R].$$

Suppose $i : X \subset Y$ finite metric spaces, and $R > d(x, y)$ for all pairs of points $x, y \in Y$ (and X).

i induces map of simplicial sheaves $V_s(X) \to V_s(Y)$, $s \in [0, R]$.

Fact: This map is a stalkwise weak equiv. if and only if $X = Y$, because $V_s(X) = X$ and $V_s(Y) = Y$ for small s.

Local homotopy theory is **not useful** (2019).

ep-metric spaces (following Spivak (2009))

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

$d_Y(f(x), f(y)) \leq d_X(x, y)$ (compresses distance, "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), $0 \leq i \leq k$, 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 polygonal 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.

Example: Say $x \sim y$ if $d(x, y) = 0$. Collapse X by equiv relation.

$ep - \mathbf{Met}$ is cocomplete

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 **coproduct** in $ep - \mathbf{Met}$.

2) Suppose given morphisms $f, g : (X, d_X) \to (Y, d_Y)$ in $ep - \mathbf{Met}$. Form the set theoretic coequalizer

$$X \underset{g}{\overset{f}{\rightrightarrows}} Y \xrightarrow{p} Z,$$

Then p is a surjective function, and we give Z the quotient ep-metric D.

$$(X, d_X) \underset{g}{\overset{f}{\rightrightarrows}} (Y, d_Y) \xrightarrow{p} (Z, D)$$

is a **coequalizer** in $ep - \mathbf{Met}$.

Vietoris-Rips complex

(X, d_X) a finite ep-metric space, $d_X : X \times X \to [0, \infty]$.

If X totally ordered (has a listing), then $V_s(X)$ has n-simplices $x_0 \leq x_1 \leq \cdots \leq x_n$ with $d_X(x_i, x_j) \leq s$ for all i, j.

$V(X) : s \mapsto V_s(X)$, $s \in [0, \infty]$ is **Vietoris-Rips system** for X.

A different way: $P_s(X)$ is the poset of all subsets $\sigma \subset X$ such that $d_X(x, y) \leq s$ for all $x, y \in \sigma$.

$P_s(X)$ is the poset of non-degenerate simplices of $V_s(X)$.

$BP_s(X)$ is the **barycentric subdivision** of $V_s(X)$. There is a natural weak equivalence $\gamma : BP_s(X) \to V_s(X)$ and a corresponding weak equivalence of systems $BP(X) \to V(X)$.

1) Poset construction $BP_s(X)$ **does not** use an ordering on X.

2) $V(X)$ and $BP(X)$ are simplicial fuzzy sets over $[0, \infty]^{op}$.

Stability

Theorem 5 (Rips stability: Blumberg-Lesnick, Memoli).

Suppose $X \subset Y$ finite metric spaces, such that $d_H(X, Y) < r$. There is a homotopy commutative diagram (homotopy interleaving)

$$\begin{array}{ccc} P_s(X) & \xrightarrow{\sigma} & P_{s+2r}(X) \\ i\downarrow & \overset{\theta}{\nearrow} & \downarrow i \\ P_s(Y) & \xrightarrow[\sigma]{} & P_{s+2r}(Y) \end{array}$$

Corollary 6 (Stability for persistence invariants).

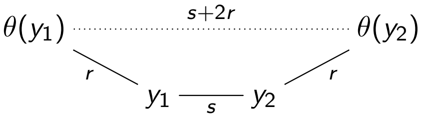
Same assumptions as Theorem 1. There are commutative diagrams

$$\begin{array}{ccc} H_k(V_s(X)) & \xrightarrow{\sigma} & H_k(V_{s+2r}(X)) \\ i\downarrow & \overset{\theta}{\nearrow} & \downarrow i \\ H_k(V_s(Y)) & \xrightarrow[\sigma]{} & H_k(V_{s+2r}(Y)) \end{array}$$

There is a corresponding statement for π_0 (clusters).

Sketch proof

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



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

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

and there are homotopies (natural transformations)

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

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

Rips stability II

Suppose X, Y are **finite metric spaces** and that an inclusion $i : X \subset Y$ defines a morphism of ep-metric spaces: $d_Y(i(x), i(y)) \leq d_X(x, y)$.

$$\textbf{compression factor}: \quad m(i) = \max_{x,y \in X} \frac{d_X(x,y)}{d_Y(i(x), i(y))}.$$

Theorem 7.

Suppose that for all $y \in Y$ there is an $x \in X$ such that $d_Y(y, i(x)) < r$. There is a homotopy commutative diagram

$$\begin{array}{ccc} P_s(X) & \xrightarrow{\sigma} & P_{m(i)(s+2r)}(X) \\ i \downarrow & \overset{\theta}{\nearrow} & \downarrow i \\ P_s(Y) & \xrightarrow[\sigma]{} & P_{m(i)(s+2r)}(Y) \end{array}$$

Applications: Want r small, $m(i)$ close to 1.

Realization (Spivak)

$X : [0, \infty] \to s\mathbf{Set}$: a **simplex** of X is a morphism $\sigma : \Delta^n \to X_s$.

A **morphism** of simplices is a commutative diagram

$$\begin{array}{ccc} \Delta^m & \xrightarrow{\theta} & \Delta^n \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\sigma} \\ X_t & \longleftarrow & X_s \end{array}$$

$s \in [0, \infty]$: U^n_s is metric on $\{0, 1, \dots, n\}$ with $d(i, j) = s$.

If $\theta : \mathbf{m} \to \mathbf{n}$ is a poset map and $s \leq t$, then θ induces an ep-metric space map $\theta : U^m_t \to U^n_s$.

$(\sigma : \Delta^n \to X_s) \mapsto U^n_s$ defines a functor $\mathbf{\Delta}/X \to ep - \mathbf{Met}$. Then

$$\mathbf{Re}(X) := \varinjlim_{\Delta^n \to X_s} U^n_s$$

is the **realization** of X.

Singular functor S

Re has a right adjoint S, called the **singular functor**.

(X, d) ep-metric space: $S(X)_{s,n} =$ sequences $(x_0, x_1, \dots, x_n)$ with $d(x_i, x_j) \leq s$ for all i, j. ("bags of words") — piece of the nerve of the trivial groupoid on X.

The adjunction map for $V(X)$ has the form

$$\eta : V(X) \to S(X),$$

with $(x_0 \leq x_1 \leq \dots \leq x_n) \mapsto (x_0, x_1, \dots, x_n)$ (forgets the ordering).

Theorem 8.

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

$$\eta : V(X)_s \to S(X)_s$$

is a weak equivalence of simplicial sets.

Proof uses simplicial approximation techniques. Show that $BNV(X)_s \to BNS(X)_s$, $\pi : \mathrm{sd}\, S(X)_s \to BNS(X)_s$ are weak equivs.

UMAP algorithm (Healy-McInnes)

X = finite set.

1) Choose **neighbourhood set** N_x, $x \in X$. Set $U_x = \{x\} \sqcup N_x$.

2) Set $(U_x, D_x) = \vee_{y \in N_x} (\{x, y\}, d_y)$ in $ep - \mathbf{Met}$. $d_y(x, y) > 0$ is a **weight**.

3) 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.

4) We have inclusions $X \subset V(X, D_x)$, $x \in X$. Form iterated pushout

$$V(X, N) = \vee_{x \in X} V(X, D_x) \simeq \vee_X S(X, D_x).$$

The diagram $V(X, N)$ is "the" **UMAP complex**.

5) Apply TDA machinery (e.g. π_0) to $V(X, N)$.

e.g. $N_x = k$ nearest neighbours if X is totally ordered, has metric.

Comparisons

(X, d) is a finite totally ordered ep-metric space, with neighbourhoods $N = \{N_x,\ x \in X\}$. If $d_x(x,y) = d(x,y)$ for $y \in N_x$, $x \in X$, there is a canonical map

$$\phi : V(X,N) = \vee_{x \in X}\ V(X, D_x) \to V(X)$$

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

Lemma 9.

$$\phi_* : \pi_0 V(X,N)_s \to \pi_0 V(X)_s$$

is a bijection if all 1*-simplices of* $V(X)_s$ *are nbhd pairs.*

Example: $N_x = k$-nearest neighbours, $r_x = \max_{y \in N_x} d(x,y)$, $s < r_x$ for all x.

Comparisons II

Fact: $V(X,N)_\infty$ is a big wedge of circles (connected).

$V(X,d_X)=\Delta^X=\Delta^M$ for $M+1=|X|$, so $V(X,N)_\infty=\vee_M\,\Delta^M$ ($M+1$ summands).

Define $\mathbf{M}\to\Delta^M=X_i$, $0\le i\le k$, $Y=\vee_N\,X_i$ (iterated pushout).

Each X_i is contractible, so $Y/X_0\simeq Y$, and

$$Y/X_0=(X_1/\mathbf{M})\vee\dots\vee(X_k/\mathbf{M})=(\Delta^M/\mathbf{M})\vee\dots\vee(\Delta^M/\mathbf{M})$$

and each

$$\begin{aligned}\Delta^M/\mathbf{M}&\simeq\Sigma\mathbf{M}\simeq\Sigma(S^0\vee\dots\vee S^0)\ \ (M\text{ summands, }\mathbf{M}\text{ pointed by }0)\\&\simeq S^1\vee\dots\vee S^1.\end{aligned}$$

Consequence: $V(X,N)_\infty\simeq\vee_{i=1}^{M^2}\,S^1\ \ (M=|X|-1)$.

Global components

Given x, y in a finite ep-metric space (X, d), say that x, y are in the same **global component** if $d(x, y) < \infty$.

Comparisons: $i : X \subset Y$ inclusion of finite sets.

Given neighbourhood sets $N_x, x \in X$, $N'_y, y \in Y$.

Suppose that $N_x \subset N'_{i(x)}$, $d_x(x, y) \geq d_{i(x)}(i(x), i(y))$ for all $y \in N_x$.

$(X, D) := \vee_{x \in X} (X, D_x)$, also (Y, D'), in $ep - \mathbf{Met}$.

Have induced map $i : (X, D) \to (Y, D')$.

If E is a global component of (X, D) then $i(E) \subset F$ for some global component F of (Y, D').

Induced map $i : (E, D) \to (F, D')$ has a compression factor $m(i)$.

Stability for UMAP

Theorem 10.

Suppose for all $y \in F$ there is an $x \in E$ such that $D'(y, i(x)) < r$. Then there is a homotopy interleaving

$$\begin{array}{ccc} P_s(E) & \xrightarrow{\sigma} & P_{m(i)\cdot(s+2r)}(E) \\ i \downarrow & \overset{\theta}{\nearrow} & \downarrow i \\ P_s(F) & \xrightarrow[\sigma]{} & P_{m(i)\cdot(s+2r)}(F) \end{array}$$

Theorem 10 follows from Theorem 7.

Lemma 11 (Excision for π_0).

$V(X, N) \to V(X, D)$ induces a bijection

$$\pi_0 V(X, N)_s \xrightarrow{\cong} \pi_0 V(X, D)_s, \ \ s \geq 0.$$

Theorem 10 + Lemma 11: stability for clusters in the global components of $V(X, N)$.

References I

Michael Barr.
Fuzzy set theory and topos theory.
Canad. Math. Bull., 29(4):501–508, 1986.

Andrew J. Blumberg and Michael Lesnick.
Universality of the homotopy interleaving distance.
CoRR, abs/1705.01690, 2017.

J.F. Jardine.
Fuzzy sets and presheaves.
Compositionality, 1:3, December 2019.

J.F. Jardine.
Metric spaces and homotopy types.
Preprint, `http://uwo.ca/math/faculty/jardine/`, 2020.

References II

Leland McInnes and John Healy.
UMAP: uniform manifold approximation and projection for dimension reduction.
CoRR, abs/1802.03426, 2018.

D.I. Spivak.
Metric realization of fuzzy simplicial sets.
Preprint, 2009.