

Fuzzy Sets and Fuzzy Techniques

Lecture 7 – Distances on and between fuzzy sets

Joakim Lindblad
joakim@cb.uu.se

Centre for Image Analysis
Uppsala University

2010-03-01

Joakim Lindblad, 2010-03-01 (1/36)

Topics of today

- Set to set distances
- (Point to set distances)
- Point to point distances

A mix of notions

- The objects that the distance is measured between (start and stop)
 - crisp or fuzzy, point or set
- The space where a path between start and stop is embedded (spatial cost function)
 - Unconstrained (Euclidean)
 - Constrained (geodesic/cost function)
- Output: Crisp (number) or fuzzy

Joakim Lindblad, 2010-03-01 (2/36)

Distances

Distances and norms are important in very many situations:

- Comparison of objects, assigning membership by examples
- Interpolation, Approximation, Measures of fulfillment of criteria
- Optimization, shortest paths, straightness
- Classification and clustering
- Neighbourhoods

A good overview:

I. Bloch. On fuzzy distances and their use in image processing under imprecision. Pattern Recognition, 32:1873–1895, 1999.

Joakim Lindblad, 2010-03-01 (3/36)

Definitions

Definition (Metric)

A **metric** is a positive function $d : X \rightarrow \mathbb{R}^+$ such that

- ① $d(x, x) = 0$ (reflexivity)
- ② $d(x, y) = 0 \Rightarrow x = y$ (separability)
- ③ $d(x, y) = d(y, x)$ (symmetry)
- ④ $d(x, z) \leq d(x, y) + d(y, z)$ (triangular inequality)

- If we drop requirement 2, we have a **pseudometric**
- If we drop requirement 4, we have a **semimetric**

A *semi-pseudometric* thus satisfies only 1 and 3.

Joakim Lindblad, 2010-03-01 (4/36)

Definitions

Definition (Norm)

A **norm** is a positive function $p : V \rightarrow \mathbb{R}^+$ such that

- ① $p(x) = 0 \Leftrightarrow x = \mathbf{0}$ (positive definiteness)
- ② $p(\lambda x) = |\lambda|p(x)$ (positive **homogeneity**)
- ③ $p(x + y) \leq p(x) + p(y)$ (triangular inequality or subadditivity)

where V is a vector space and $\lambda \in \mathbb{R}$.

Any normed vector space $(V, \|\cdot\|)$ is a metric space under the metric $d : V \times V \rightarrow \mathbb{R}$ given by $d(u, v) = \|u - v\|$

Moving to fuzzy

Approaches

- generalize crisp definitions
- infer distance from similarity (often shape similarity of sets)
- deduce distance from set relations (or other relations)
- symbolic (language/graphs)
- ...

Notions

- Distance relations
- Similarity relations ($S : \mathcal{F} \times \mathcal{F} \rightarrow [0, 1]$)
- Proximity relations (reflex. and symm.)
- Similitude, satisfiability, inclusion, resemblance, dissimilarity, ...

Set to set distances

Distances between fuzzy sets

- a) Membership focused (vertical, co-domain)
- b) Spatially focused (horizontal, domain)
- c) Mix of spatial and membership (tolerance)
- d) Feature distances (low or high dimensional representations)
- e) Morphological (mixed focus)

Membership focused

L_p norm

“The functional approach”

The most common:

Based on the family of **Minkowski distances**

$$d_p(A, B) = \left(\int_X |\mu_A(x) - \mu_B(x)|^p dx \right)^{1/p}, \quad p \geq 1,$$

$$d_{\text{EssSup}}(A, B) = \lim_{p \rightarrow \infty} d_p(A, B)$$

$$d_\infty(A, B) = \sup_{x \in X} |\mu_A(x) - \mu_B(x)|.$$

d_p is a pseudometric in the continuous case, (e.g. $d(\mathbb{Q}, \mathbb{Z}) = 0$). d_∞ is a true metric.

Membership focused

 L_p norm

Discrete version:

$$d_p(A, B) = \left(\sum_{i=1}^n |\mu_A(x_i) - \mu_B(x_i)|^p \right)^{1/p}, \quad p \geq 1,$$

$$d_\infty(A, B) = \max_{i=1, \dots, n} (|\mu_A(x_i) - \mu_B(x_i)|).$$

d_p for $p \geq 1$ are all metrics in the discrete case. If $p < 1$ we lose triangular inequality (i.e., semimetric).

Normalized variants, divide with $|X|$, or $|A| + |B|$, or similar.

Joakim Lindblad, 2010-03-01 (9/36)

Membership focused

Set operations approach

Based on this and several other assumptions, Tversky derived the following relationship:

$$S(A, B) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$$

Here, S is an interval scale of similarity, f is an interval scale that reflects the salience of the various features (typically cardinality $|\cdot|$), and θ, α and β are parameters that provide for differences in focus on the different components.

Relative version:

$$S'(A, B) = \frac{f(A \cap B)}{[f(A \cap B) + \alpha f(A \cap \bar{B}) + \beta f(B \cap \bar{A})]}$$

Joakim Lindblad, 2010-03-01 (11/36)

Membership focused

Set operations approach

Tversky 1977, et al. A model based on identifying three sets; what is unique to A , to B , and what is common to A and B .

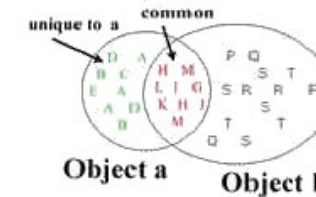


Figure: Representation of two objects that each contains its own unique features and also contains common features.

An important aspect of Tversky's model is that similarity depends not only on the proportion of features common to the two objects but also on their unique features. Each letter here represents a feature.

Joakim Lindblad, 2010-03-01 (10/36)

Similarity measures

Jaccard/Tanimoto

$$S(\vec{A}, \vec{B}) = \frac{A \cap B}{A \cup B} \quad (1)$$

Tversky

$$S(\vec{A}, \vec{B}) = \frac{A \cap B}{\alpha A + \beta B + (1 - \alpha - \beta)(A \cap B)} \quad (2)$$

Baroni-Urbani/Buser

$$S(\vec{A}, \vec{B}) = \frac{\sqrt{A \cap B \cdot (N - A \cup B)} + A \cap B}{\sqrt{A \cap B \cdot (N - A \cup B)} + A \cup B} \quad (3)$$

Ochiai/Cosine

$$S(\vec{A}, \vec{B}) = \frac{A \cap B}{\sqrt{AB}} \quad (4)$$

Pearson

$$S(\vec{A}, \vec{B}) = \frac{N(A \cap B) - AB}{\sqrt{AB(N - A)(N - B)}} \quad (5)$$

Stiles

$$S(\vec{A}, \vec{B}) = \log_{10} \frac{N \cdot (N(A \cap B) - AB) - \frac{A^2}{2}}{AB(N - A)(N - B)} \quad (6)$$

McConnaughey

$$S(\vec{A}, \vec{B}) = \frac{(A \cap B)(2 - A - B) + AB}{AB} \quad (7)$$

Dennis measure

$$S(\vec{A}, \vec{B}) = \frac{N(A \cap B) - AB}{\sqrt{NAB}} \quad (8)$$

Russel/Rao

$$S(\vec{A}, \vec{B}) = \frac{A \cap B}{N} \quad (9)$$

Dice

$$S(\vec{A}, \vec{B}) = \frac{2(A \cap B)}{A + B} \quad (10)$$

Sokal/Sneath(1)

$$S(\vec{A}, \vec{B}) = \frac{A \cap B}{2A + 2B + A \cap B} \quad (11)$$

Sokal/Sneath(2)

$$S(\vec{A}, \vec{B}) = \frac{2 \cdot (N - A - B + 2(A \cap B))}{2 \cdot N - A - B + 2(A \cap B)} \quad (12)$$

Sokal/Sneath(3)

$$S(\vec{A}, \vec{B}) = \frac{N - A - B + 2(A \cap B)}{A + B - 2(A \cap B)} \quad (13)$$

Joakim Lindblad, 2010-03-01 (12/36)

Similarity measures

Simple Matching

$$S(\tilde{A}, \tilde{B}) = \frac{N - A - B + 2(A \cap B)}{N} \quad (14)$$

Kulczynski(1)

$$S(\tilde{A}, \tilde{B}) = \frac{A \cap B}{A + B - 2(A \cap B)} \quad (15)$$

Kulczynski(2)

$$S(\tilde{A}, \tilde{B}) = \frac{(A \cap B)(A + B)}{2AB} \quad (16)$$

Hamann

$$S(\tilde{A}, \tilde{B}) = \frac{N + 4(A \cap B) - 2A - 2B}{N} \quad (17)$$

Rogers/Thaimoto

$$S(\tilde{A}, \tilde{B}) = \frac{N - A - B + 2(A \cap B)}{N + A + B - 2(A \cap B)} \quad (18)$$

Forbes

$$S(\tilde{A}, \tilde{B}) = \frac{N(A \cap B)}{AB} \quad (19)$$

Yule

$$S(\tilde{A}, \tilde{B}) = \frac{N(A \cap B) - AB}{(A \cap B)(N - 2A - 2B + 2) + AB} \quad (20)$$

Fossum

$$S(\tilde{A}, \tilde{B}) = \frac{N((A \cap B) - 0.5)^2}{AB} \quad (21)$$

Simpson

$$S(\tilde{A}, \tilde{B}) = \frac{A \cap B}{\min(A, B)} \quad (22)$$

Mean Manhattan

$$S(\tilde{A}, \tilde{B}) = \frac{A + B - 2(A \cap B)}{N} \quad (23)$$

Normalized Euclidian

$$S(\tilde{A}, \tilde{B}) = \sqrt{\frac{A + B - 2(A \cap B)}{N}} \quad (24)$$

Joakim Lindblad, 2010-03-01 (13/36)

Spatially focused

- Nearest point
- Mean distance
- Hausdorff

Three (four) approaches:

- fuzzy distance
- weighting with membership
- morphological and integration of alpha-cuts

Joakim Lindblad, 2010-03-01 (14/36)

Spatially focused

A fuzzy shortest distance between sets

Dubois and Prade (1980)

$$\bar{d}_{A,B}(r) = \sup_{\substack{x,y \\ d(x,y)=r}} [\inf(A(x), B(y))]$$

Is **not** a generalization of *shortest* distance between two crisp sets; rather the *set of distances* between two sets.

Rosenfeld (1985)

Distance distribution

$$\Delta_{A,B}(r) = \sup_{\substack{x,y \\ d(x,y) \leq r}} [\inf(A(x), B(y))]$$

Distance density

$$\int_{-\infty}^r \delta_{A,B}(x) dx = \Delta_{A,B}(r)$$

Joakim Lindblad, 2010-03-01 (15/36)

Spatially focused

Mean distance

Mean distance

Crisp:

$$\frac{\sum_{x \in A} \sum_{y \in B} d(x, y)}{|A||B|}$$

Fuzzy (Rosenfeld 1985):

$$\frac{\sum_{x \in A} \sum_{y \in B} d(x, y) \inf[A(x), B(y)]}{\sum_{x \in A} \sum_{y \in B} \inf[A(x), B(y)]}$$

No metric: $d(A, A) > 0$ in general.

Joakim Lindblad, 2010-03-01 (16/36)

Spatially focused

Hausdorff

Crisp:

$$d_H(A, B) = \max\left\{\sup_{x \in A} \inf_{y \in B} d(x, y), \sup_{y \in B} \inf_{x \in A} d(x, y)\right\}$$

$$= \inf\{r \in \mathbb{R}^+ \mid A \subseteq D_r(B) \wedge B \subseteq D_r(A)\}$$

where $D_r(A)$ is the dilation of the set A by a ball of radius r

$$D_r(A) = \{y \in X \mid \exists x \in A : d(x, y) \leq r\}$$

The Hausdorff distance between A and B is the smallest amount that A must be expanded to contain B and vice versa.

Is a metric on the set of nonempty compact sets.

Remark:

Usually extended with: $d_H(A, \emptyset) = \infty$ and $d_H(\emptyset, \emptyset) = 0$

Joakim Lindblad, 2010-03-01 (17/36)

Spatially focused

Hausdorff

Ralescu and Ralescu (1984)

$$d_{H_1}(A, B) = \int_0^1 d_H({}^\alpha A, {}^\alpha B) d\alpha,$$

$$d_{H_\infty}(A, B) = \sup_{\alpha > 0} d_H({}^\alpha A, {}^\alpha B),$$

where d_H is the Hausdorff distance between two crisp sets,

Again, a serious problem is that the distance between two fuzzy sets A and B is infinite if $\text{height}(A) \neq \text{height}(B)$.

A number of variations have been suggested to overcome this problem. . .

Joakim Lindblad, 2010-03-01 (19/36)

Spatially focused

Hausdorff

Fuzzy:

Dubois and Prade (1980):

Define dilation with a local max operator

$$D_r(A)(x) = \sup_{y \in X, d(x,y) \leq r} A(y)$$

$$d_H(A, B) = \inf\{r \in \mathbb{R}^+ \mid A \subseteq D_r(B) \wedge B \subseteq D_r(A)\}$$

However, since the dilation is only "horizontal", we may never be able to cover the other set. That is, if A and B have different *height*.

Joakim Lindblad, 2010-03-01 (18/36)

Spatially focused

Hausdorff: Attacking the problem of different heights. . .

Chaudhuri and Rosenfeld (1996):

Modifying the sets

$$d_H(A, B) = \int_0^1 w(\alpha) d_H({}^\alpha A', {}^\alpha B') d\alpha + \varepsilon \frac{d_1(A, B)}{|X|},$$

where $w(\alpha)$ may be any function $\int_0^1 w(\alpha) d\alpha = 1$, A' and B' are normal fuzzy sets, such that $A'(x) = A(x)$ where $A(x) < \text{height}(A)$ and $A'(x) = 1$ otherwise, d_1 is the L_1 -norm distance, and ε is a small (ugly) value.

Joakim Lindblad, 2010-03-01 (20/36)

Spatially focused

Hausdorff: Attacking the problem of different heights...

Boxer (1997):

Using the complements of the α -cuts.

$$d_H(A, B) = \int_0^1 w(\alpha) d_H({}^\alpha \bar{A}, {}^\alpha \bar{B}) d\alpha$$

Not consistent with crisp Hausdorff distance.

Joakim Lindblad, 2010-03-01 (21/36)

Spatially focused

Hausdorff: Attacking the problem of different heights...

Fan (1997):

Limit the distance to $c_{\max}$

$$d_H(A, B) = \int_0^1 w(\alpha) \min [c_{\max}, d_H({}^\alpha A, {}^\alpha B)] d\alpha$$

Joakim Lindblad, 2010-03-01 (22/36)

Spatially focused

Hausdorff: Finally!

Brass (2002):

proves that it is **not possible** to construct an “intuitive” Hausdorff distance on the set of fuzzy sets.

That is, satisfying the following properties:

- Metric
- Translation invariant
- Independent on scale of underlying set (length unit shouldn't matter)

Joakim Lindblad, 2010-03-01 (23/36)

Mixed focus with spatial tolerance

Lowen and Peeters (1998)

Spatial tolerance combined with membership based distance

Introduce the local difference at a point x

$$d_x^t(A, B) = \inf_{\substack{d(x, y) \leq t \\ d(x, z) \leq t}} |A(y) - B(z)|$$

in the Minkowski distance, $p \geq 1$

$$d_p^t(A, B) = \left(\int_X (d_x^t(A, B))^p \right)^{1/p}$$

d_p^t is a semi-pseudometric, d_p^t is decreasing as $t \rightarrow 0$, for Riemann-integrable sets, $\lim_{p \rightarrow 0} d_p^t(A, B) \rightarrow d_p(A, B)$.

Joakim Lindblad, 2010-03-01 (24/36)

Feature distances

"Pattern recognition approach"

Use of a **feature representation** of the observed sets as an intermediate step in the distance calculations.

The distance between sets A and B is then given in terms of the distance between their feature vectors.

Often global shape features are used (think shape matching, image retrieval).

Feature distances

Combining local and global features

Relying only on global information in general provides a very rough pseudo-metric, where a lot of different sets have distance zero to each other.

By using a high dimensional feature representation, capturing the local structure of the sets (possibly by including the membership of all individual points), separability can be achieved.

Feature distances

Given an injective function Φ from $\mathcal{F}(X)$ to a metric space H , we can define a metric on $\mathcal{F}(X)$ by requiring that Φ is an isometry.

That is, the distance between fuzzy sets A and B is

$$d^\Phi(A, B) = d(\Phi(A), \Phi(B)) . \quad (1)$$

The metric in the feature space H is often given by a Minkowski type distance measure.

Morphological approach

Starting with a **point to set distance**

Morphological definition of a crisp point to set distance:

$$\begin{aligned} d_B(x, X) &= 0 \Leftrightarrow x \in X \\ d_B(x, X) &= r \Leftrightarrow x \in D_r(X) \quad \text{and} \quad x \notin D_{r-\varepsilon}(X) \end{aligned}$$

where D_r denotes dilation with a ball or radius r .

Generalizing all notions to the fuzzy case we get the following **fuzzy** point to set distance:

$$\begin{aligned} \delta(x, X)(0) &= X(x) \\ \delta(x, X)(r) &= \min [D_r(X)(x), 1 - D_{r-\varepsilon}(X)(x)] \end{aligned}$$

Morphological approach

From that we define a nearest point set to set distance

$$d(X, Y) = n \Leftrightarrow D_n(X) \cap Y \neq \emptyset \quad \text{and} \quad D_{n-\varepsilon}(X) \cap Y = \emptyset$$

similarly in the fuzzy case

$$d_N(X, Y)(n) = \min \left[\max_z \{ \min(Y(z), D_n(X)(z)) \}, \right. \\ \left. 1 - \max_z \{ \min(Y(z), D_{n-\varepsilon}(X)(z)) \} \right]$$

or rather the symmetrical one! (like Hausdorff)

Comparable with Rosenfelds distance distribution.

Alternatives

Graph approach – see feature based approach

Cost of deforming one set into another (earth movers distance)

In terms of linguistic variables, near, far, etc.

Point to point distances

Distances between points in a fuzzy set

Defining the cost of traveling along a path

Crisp case:

- a) Euclidean or other L_p metrics
- b) Geodesic/constrained (crisp constraints)
- c) Cost function (grey-level image)

Fuzzy case:

- a) Same
- b) Fuzzy constraints
- c) Similar

Cost function

snow shoveling distance

Similar to grey weighted distances (Rutovitz '68, Levi & Montanari '70) put in a fuzzy framework (Saha '02).

Define the distance along a path π_i between points x and y in the fuzzy set A

$$d_A(\pi_i(x, y)) = \int_{s \in \pi} A(t) dt$$

The distance between points x and y in A is the distance along the shortest path

$$d_A(x, y) = \inf_{\pi \in \Pi(x, y)} d_A(\pi)$$

out of all possible paths between x and y , $\Pi(x, y)$.

Cost function

variations

Membership as another dimension
integrate the arc-length

Bloch 1995, Toivanen 1996:

$$d_A(\pi) = \int_{s \in \pi} \sqrt{1 + \left(\frac{dA(t)}{dt} \right)^2} dt$$

Problem: Scale of membership relative to spatial distance

Constrained distance

Geodesic distance – shortest path within the set; not allowed to
go out of the set

a path that descends the least in terms of membership

Constrained distance

Connectedness, Rosenfeld 1979

Strength of a path – the strength of its weakest link

Strength of a link between two points defined by the
membership function.

The connectedness of two points x and y in A –
the strength of the strongest path between x and y

$$c_A(x, y) = \sup_{\pi \in \Pi(x, y)} \inf_{t \in \pi} A(t)$$

Constrained distance

Bloch and Maître 1995:

$$d(x, y) = \inf_{\pi \in \Pi_{c_A}(x, y)} \frac{\int_{\pi} ds}{c_A(x, y)}$$

where $c_A(x, y)$ is the strength of connectedness of points x and
 y , and $\Pi_{c_A}(x, y)$ is the set of path contained within the α -cut
 $c_A A$.

Does not provide triangle inequality.