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

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:

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

Definitions
Definition (Metric)
A metric is a positive function $d : X \rightarrow \mathbb{R}^+$ such that
1. $d(x, x) = 0$ (reflexivity)
2. $d(x, y) = 0 \implies x = y$ (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.
Definitions

Definition (Norm)

A norm is a positive function \( p : V \to \mathbb{R}^+ \) such that

1. \( p(x) = 0 \iff x = 0 \) (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 \( d : V \times V \to \mathbb{R} \) given by \( d(u, v) = \|u - v\| \).

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

Membership focused

\( L_p \) norm

"The functional approach"

The most common:

Based on the family of Minkowski distances

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

\[
d_{\text{EssSup}}(A, B) = \lim_{p \to \infty} d_p(A, B)
\]

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

\( d_p \) is a pseudometric in the continuous case, (e.g. \( d(Q, Z) = 0 \)). \( d_\infty \) is a true metric.
Membership focused

$L_p$ norm

Discrete version:

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

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

$d_p$ 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.

---

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 $\theta, \alpha$ and $\beta$ are parameters that provide for differences in focus on the different components.

Relative version:

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

---

Similarity measures

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 $B$.

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.
Similarity measures

<table>
<thead>
<tr>
<th>Fuzzy Sets and Fuzzy Techniques</th>
<th>Joakim Lindblad</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outline</td>
<td>Distances</td>
</tr>
<tr>
<td>Set to set distances</td>
<td>Membership focused</td>
</tr>
<tr>
<td>Spatially focused</td>
<td>Mixed focus with spatial tolerance</td>
</tr>
<tr>
<td>Feature distances</td>
<td>Morphological approach</td>
</tr>
<tr>
<td>Point to point distances</td>
<td></td>
</tr>
</tbody>
</table>

### Spatially focused

A fuzzy shortest distance between sets

Dubois and Prade (1980)

\[
\bar{d}_{A,B}(r) = \sup_{x,y \in A \times B} \left[ \inf(A(x), B(y)) \right] \quad (14)
\]

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

Rosenfeld (1985)

Distance distribution

\[
\Delta_{A,B}(r) = \sup_{x,y \in A \times B} \left[ \inf(A(x), B(y)) \right] \quad (15)
\]

Distance density

\[
\int_{-\infty}^{r} \delta_{A,B}(x) \, dx = \Delta_{A,B}(r) \quad (16)
\]

### Spatially focused

- Nearest point
- Mean distance
- Hausdorff

Three (four) approaches:
- fuzzy distance
- weighting with membership
- morphological and integration of alpha-cuts

### Spatially focused

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.
is a small (ugly) value. ε

\[ d_1 = \inf \{ r \in \mathbb{R}^+ \mid A \subseteq D_r(B) \land B \subseteq D_r(A) \} \]

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 \( d_H(\emptyset, \emptyset) = 0 \)

Ralescu and Ralescu (1984)

\[
d_{H_1}(A, B) = \int_0^1 d_H(^0A, ^\alpha B) \, d\alpha, \\
d_{H_\infty}(A, B) = \sup_{\alpha > 0} d_H(^0A, ^\alpha B),
\]

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...

Fuzzy:

 Dubois and Prade (1980):

Define dilation with a local max operator

\[
D_r(A)(x) = \sup_{y \in X} A(y) \quad \text{if} \quad d(x, y) \leq r
\]

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

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.

Ralescu and Ralescu (1984)

\[
d_{H_1}(A, B) = \int_0^1 d_H(^0A, ^\alpha B) \, d\alpha, \\
d_{H_\infty}(A, B) = \sup_{\alpha > 0} d_H(^0A, ^\alpha B),
\]

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...

Chaudhuri and Rosenfeld (1996):

Modifying the sets

\[
d_H(A, B) = \int_0^1 w(\alpha) d_H(^0A', ^\alpha B') \, d\alpha + \varepsilon \left( \frac{1}{|X|} \right),
\]

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 \( \varepsilon \) is a small (ugly) value.
 Spatially focused

Hausdorff: Attacking the problem of different heights...

Boxer (1997):
Using the complements of the $\alpha$-cuts.

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

Not consistent with crisp Hausdorff distance.

Fan (1997):
Limit the distance to $c_{\text{max}}$

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

Mixed focus with spatial tolerance

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)

Lowen and Peeters (1998)

Spatial tolerance combined with membership based distance
Introduce the local difference at a point $x$

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

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 $t \to 0$,
for Riemann-integrable sets, $\lim_{t \to 0} d_p^t(A, B) = d_p(A, B)$. 
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).

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

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

$$d^\Phi(A, B) = d(\Phi(A), \Phi(B)).$$  \hspace{1cm} (1)

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

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.

Starting with a point to set distance

Morphological definition of a crisp point to set distance:

$$d_B(x, X) = 0 \iff x \in X$$
$$d_B(x, X) = r \iff x \in D_r(X) \text{ and } x \notin D_{r-\varepsilon}(X)$$

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:

$$\delta(x, X)(0) = X(x)$$
$$\delta(x, X)(r) = \min [D_r(X)(x), 1 - D_{r-\varepsilon}(X)(x)]$$
Morphological approach

From that we define a nearest point set to set distance

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

similarly in the fuzzy case

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

or rather the symmetrical one! (like Hausdorff)
Comparable with Rosenfelds distance distribution.

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 \( \pi \) between points \( x \) and \( y \) in the fuzzy set \( A \)

\[ d_A(\pi_i(x, y)) = \int_{\pi} A(t) \, dt \]

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 \( y \), \( \Pi(x, y) \).
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

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 \( y \)

\[ c_A(x, y) = \sup_{\pi \in \Pi_{cA}(x, y)} \inf_{t \in \pi} A(t) \]

Bloch and Maître 1995:

\[ d(x, y) = \inf_{\pi \in \Pi_{cA}(x, y)} \int_{\pi} ds \]

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

Does not provide triangle inequality.