FUZZY DIFFERENTIAL GEOMETRY - THEORY AND APPLICATIONS TO MACHINE LEARNING AND MANIFOLD ESTIMATION

YOUSUF M. SOLIMAN

ABSTRACT. I explored this topic to get a better feel for fuzzy logic. Nothing here should be taken too seriously...

1. INTRODUCTION

Historically, the notion of a differentiable manifold, that is, a set that looks locally like Euclidean Space, has been an integral part of various fields of mathematics. One may note their applications in the fields of Differential Topology [1], Algebraic Geometry, Algebraic Topology, and Lie Groups and their associated algebras. We will base our work upon the already well-established fuzzy structures viz. fuzzy topological spaces, fuzzy topological vector spaces, fuzzy derivatives. However, the definition of a fuzzy derivative provided in Foster [1] is not easily generalized to a general k derivatives. Consequently, the existence of a fuzzy differentiable manifold of class greater than one has not yet been established. We shall give topological separation axioms that have not been given previously, for sake of completeness.

Our principal approach is quite similar to the methods used in [1]. Namely, we shall take definitions in [2] and [3] of fuzzy continuity and fuzzy topological vector space, and use these notions to give a fuzzy topological vector space differential structure by constructing a fuzzy homeomorphism and, naturally, a fuzzy diffeomorphism of class k. To do so, we provide a new definition of a fuzzy object, known as fuzzy vectors. We then define proper fuzzy directional derivatives along these abstract fuzzy vectors, and allude to their applications in manifold learning. For completeness, we shall define tangent vector spaces to these fuzzy manifolds.

In regards to the notion of a fuzzy vector, we will define a fuzzy $L^p$ norm that reduces analogously. We shall also define such analysis notions as fuzzy Banach spaces, as well as a natural definition of fuzzy Lebesgue integration. We shall give criteria for integrability analogous to those given for non-fuzzy mappings.

After we properly establish the notion of a fuzzy differentiable manifold of class k, we shall consider operations on these objects. Namely, we may first discuss maps between these manifolds, as well as their associated fuzzy tangent spaces. We then glue these tangent spaces together to obtain a fuzzy tangent bundle. We have the notion of a covariant functor between $\mathcal{C}^{r+1}$ and $\mathcal{C}^r$ fuzzy differentiable manifolds in this way. We proceed to define the fuzzy cotangent bundle, as well as fuzzy differential forms.

2. PRELIMINARIES

2.1. Fuzzy Sets and Notation on X. We assume the reader is not familiar with fuzzy sets and their associated notation; consequently, we outline such notation and operations here. However, we assume knowledge of Set Theory, Topology, Differential Topology, Real Analysis, Abstract Algebra and a general mathematical maturity. For the computer scientist, the arguments may appear too theoretical. On the contrary, these well-established foundations will allow for proper implementation.

Definition 2.1. Let $X$ be a set and $I := [0, 1]$. A fuzzy set $A$ in $X$ is a set uniquely determined by its membership function $\mu_A : x \in X \rightarrow [0, 1]$, where $\mu_A(x)$ is the point x’s “grade of membership”.

In this way, fuzzy sets are generalizations of familiar sets. That is, if $\mu_A$ takes only the values 0 and 1, then we can think of the fuzzy set A as an ordinary set whose elements are those $x$ with $\mu_A(x) = 1$. We can define the usual set-theoretic operations on fuzzy sets. All of these operations reduce to the familiar set operations when $\mu_A$ has range {0, 1}.

Definition 2.2. Let $A$ and $B$ be fuzzy sets in $X$. Then we have $A = B \iff \mu_A(x) = \mu_B(x)$, $A \subseteq B \iff \mu_A(x) \leq \mu_B(x)$, $C = A \cup B \iff \mu_C(x) = \max\{\mu_A(x), \mu_B(x)\}$, $D = A \cap B \iff \mu_D(x) = \min\{\mu_A(x), \mu_B(x)\}$, and $E = A^c \iff \mu_E(x) = 1 - \mu_A(x)$. 

---

Date: May 4, 2016.
1 − \mu_A(x).
More generally, given a family \{A_j\}_{j \in J},
\[ D = \bigcup_{j \in J} A_j, \quad C = \bigcap_{j \in J} A_j, \]
\[ \mu_D(x) = \sup_{j \in J} \mu_{A_j}(x), \quad \mu_C(x) = \inf_{j \in J} \mu_{A_j}(x). \]

**Definition 2.3.** Let \( X \) and \( Y \) be sets, and let \( f : X \to Y \) be a mapping. For a fuzzy set \( B \in Y \), the “inverse image” of \( B \) under \( f \) is the fuzzy set \( A \) with the membership function defined by
\[ \mu_A(x) := \mu_B(f(x)) \]
for all \( x \) in \( X \).

Consider now the converse. The question is “what is the membership function for the fuzzy set \( B \) in \( Y \) which is induced by \( f \)?”

If \( f \) is not injective, then ambiguity arises when two or more distinct points in \( X \), with different grades of membership in \( A \), are mapped into the same point \( y \) in \( Y \). To resolve this ambiguity, we assign the larger of the two grades of membership to \( y \). More generally, the membership function for \( B \) is:
\[ \mu_B(y) = \begin{cases} \sup_{x \in f^{-1}(y)} \mu_A(x) & : f^{-1}(y) \neq \emptyset \\ 0 & : \text{null} \end{cases} \]
where \( f^{-1}(y) = \{ x \in X \mid f(x) = y \} \)

**Definition 2.4.** A fuzzy point in \( X \) is a fuzzy set with membership function \( \mu_{y_\lambda}(x), x \in X \), defined as
\[ \mu_{y_\lambda}(x) = \begin{cases} \mu_{y_\lambda}(x) = \lambda & : x = y \\ 0 & : \text{else} \end{cases} \]
We denote by \( \mu_{y_\lambda}(x) \) the membership function such that \( \mu_{y_\lambda}(x) = c, \forall x \in X \).

**Definition 2.5.** Let \( F(\mathbb{R}) \) denote the set of all fuzzy subsets on \( \mathbb{R} \). For every \( \mu \in F(\mathbb{R}) \), \( \mu \) is a fuzzy number if \( \mu \) has the following properties:
- \( \mu \) is normal (\( \exists x_0 \in \mathbb{R} | \mu(x_0) = 1 \))
- \( \mu \) is convex (\( \mu(\lambda x + (1-\lambda)y) \geq \min\{\mu(x), \mu(y)\} \) for all \( x, y \in \mathbb{R}, \lambda \in I \))
- \( \mu \) is upper semicontinuous

We denote by \( \mathbb{E} \) the set of all fuzzy numbers. It should be noted that \( \mathbb{R} \) can be easily embedded into \( \mathbb{E} \). We say that \( \mu \in \mathbb{E} \) is non-negative if \( \mu(x) = 0 \) \( \forall x < 0 \). The set of all non-negative fuzzy numbers is denoted by \( \mathbb{E}^+ \).

### 2.2 Fuzzy Sets and Notation on \( X^n \)

We use the notation \( X^n \) to mean a product of sets in some index set \( J \) with \( \#J = n \). In \( X^n \) there are two different categories of fuzzy sets we may discuss. Namely, traditional fuzzy sets defined on \( X^n \) and fuzzy vectors induced by cartesian products.

**Definition 2.6.** Let \( X^n \) be a product of sets. Let \( A \subset X^n \). We define the fuzzy membership function as
\[ \mu_A(x) : X^n \to \prod_k I \]
with \( k \leq n \).

In most of the work, we shall deal with the case \( k = n \), for it is the most useful in application, as well as the most clean to work with. It shall be henceforth assumed that \( k = n \), unless otherwise noted.

**Definition 2.7.** Let \( \{A_j\} \) be a family of fuzzy sets on \( X \), we define a fuzzy vector as a member of the product
\[ x_\lambda \in A = \prod_{j \in J} A_j \quad \text{with} \quad x_\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \]
Define a fuzzy $L^p$ norm on a fuzzy vector as

$$||x||_p = \left( \sum_{j \in J} (y_{\lambda_j})^p \right)^{\frac{1}{p}}$$

We may say this is a norm since it inherits the same properties as a standard $L^p$ norm. That is, if $\lambda_j$ is 1, this reduces to the $L^p$ norm on $x$.

with the 0 vector having norm 0, written as $0 = \begin{pmatrix} k_0 \\ k_1 \\ \vdots \\ k_0 \end{pmatrix}$, $||0||_p = 0$.

Now we must define natural extensions of binary operations to sets of fuzzy vectors.

**Definition 2.8.** Let $A$ and $B$ be collections of fuzzy vectors in $X^n$, we say that $A = B \iff \mu_A(x) = \mu_B(x), A = B \iff \mu_A(x) \leq \mu_B(x)$

For some index set $K$,

$$C = \bigcup_{k \in K} A_k \iff \mu_C(x) = \sup_{k \in K} \{ \mu_A(x) \mid A \in A_k \} = \begin{pmatrix} \sup_{k \in K} \{ \mu_{A_{k_1}}(x_{k_1}) \} \\ \sup_{k \in K} \{ \mu_{A_{k_2}}(x_{k_2}) \} \\ \vdots \\ \sup_{k \in K} \{ \mu_{A_{k_n}}(x_{k_n}) \} \end{pmatrix}$$

For some index set $K$,

$$D = \bigcap_{k \in K} A_k \iff \mu_D(x) = \inf_{k \in K} \{ \mu_A(x) \mid A \in A_k \} = \begin{pmatrix} \inf_{k \in K} \{ \mu_{A_{k_1}}(x_{k_1}) \} \\ \inf_{k \in K} \{ \mu_{A_{k_2}}(x_{k_2}) \} \\ \vdots \\ \inf_{k \in K} \{ \mu_{A_{k_n}}(x_{k_n}) \} \end{pmatrix}$$

It is natural to define things such as the inverse image and image of a collection of fuzzy vectors in an analogous way, just the component-wise inverse images and images of each element of a fuzzy vector, just as we have done for unions and intersections. For this reason, we omit explicit definitions.

3. Fuzzy Metrics

3.1. Hausdorff Distance. Naturally, the fuzzy $L^p$ norm induces a metric on fuzzy vectors. Namely, $||x - y||_p = d(x, y)$. However, we wish to discuss metrics in a notion of fuzzy Hausdorff Distance, as such a distance was proven to be not well-defined. We shall therefore consider level sets of fuzzy sets as a means of developing a notion of a fuzzy Hausdorff distance on these sets.

Let $A$ and $B$ be two nonempty fuzzy sets in some metric space $X$. We say the maximum membership of a fuzzy set $A$ is:

$$\alpha^* = \max \{ \mu_A(x) \mid x \in X \}$$

We then define the non-fuzzy set $\mathcal{A}_{\text{max}}$ as follows:

$$\mathcal{A}_{\text{max}} = \{ x \mid \mu_A(x) = \alpha^* \}$$

Let $\mathcal{A}_t$ be a nonempty non-fuzzy subset of $X$ such that

$$\mathcal{A}_{\text{max}} \subseteq \mathcal{A}_t$$

and such that for any two fuzzy sets $A$ and $B$

$$\mathcal{A}_t = \mathcal{A}_t \iff \mathcal{A}_{\text{max}} = \mathcal{B}_{\text{max}}$$

We define the family of fuzzy sets $\mathcal{A}_t$, where $t \in [0, 1]$, by

$$\mathcal{A}_t = \begin{cases} \{ x \mid \mu_A(x) \in [t, \alpha^*] \} & \text{if } t \leq \alpha^* \\ \mathcal{A}_t & \text{if } t > \alpha^* \end{cases}$$

Not that $\mathcal{A}_t = \mathcal{A}_{\text{max}}$ if $t = \alpha^*$ and that the second case cannot arise if $\alpha^* = 1$. 


First, we will suppose that the membership functions only take on a discrete set of membership values \( t_j \) where \( j \in J \) and \( J \) is an index set. Let \( H(\mathcal{A}, \mathcal{B}) \) be the regular Hausdorff distance between \( \mathcal{A} \) and \( \mathcal{B} \). That is,

\[
H(\mathcal{A}, \mathcal{B}) = \max \{ \sup_{a \in \mathcal{A}} \inf_{b \in \mathcal{B}} d(a, b), \sup_{b \in \mathcal{B}} \inf_{a \in \mathcal{A}} d(a, b) \}
\]

We now define the fuzzy set distance to be

\[
\mathcal{H}(A, B) = \frac{\sum_{t \in J} t \cdot H(\mathcal{A}_t, \mathcal{B}_t)}{\sum_{t \in J} t}
\]

Now we will consider the case when \( A \) and \( B \) are continuous valued. Analogous to the discrete case above, we define the fuzzy set distance to be

\[
\mathcal{H}(A, B) = \int_0^1 t \cdot H(\mathcal{A}_t, \mathcal{B}_t) dt = 2 \int_0^1 t H(\mathcal{A}_t, \mathcal{B}_t) dt
\]

It is clear that our definition depends on how we choose \( \mathcal{A}_t \). If a fuzzy set \( A \) has a maximum membership value \( \alpha^* = 1 \) then \( \mathcal{A}_t \) does not need to be defined since \( \mathcal{A}_t \) does not rely on it’s definition. However, if \( \alpha^* < 1 \), we can defined \( \mathcal{A}_t \) by the union of \( \mathcal{A}_{\text{max}} \) and \( \{ x_\alpha \} \) (\( \mathcal{A}_t = \mathcal{A}_{\text{max}} \cup \{ x_\alpha \} \)). This single point may very well have a negligible effect on the distance by simply giving it negligible area. As long as we choose \( \mathcal{A}_t \) consistently our results will be consistent.

We will show that this metric is indeed Hausdorff and reduces to the regular Hausdorff metric for crisp sets.

Proposition 3.1. For any nonempty fuzzy set \( A \): \( \mathcal{A}_t \neq \emptyset \) \( \forall t \in [0, 1] \).

Proposition 3.2. If \( A = B \) then for all \( t \) we have \( \mathcal{A}_t = \mathcal{B}_t \).

Proposition 3.3. If \( A \neq B \), there exists \( t > 0 \) such that \( \mathcal{A}_t \neq \mathcal{B}_t \).

Proof. For this proof we consider two cases: (1) when \( A \) and \( B \) have the same maximum membership value, and (2) when they have different maximum membership values.

Case 1: Let \( \alpha^* = \beta^* \). Since \( A = B \), there is a point \( x_0 \in X \) such that \( \mu_A(x_0) \neq \mu_B(x_0) \). If \( \mu_A(x_0) > \mu_B(x_0) \) then, by definition we have \( \mathcal{A}_{\mu_B(x_0)} \neq \mathcal{B}_{\mu_B(x_0)} \). We can easily see that \( x_0 \in \mathcal{A}_{\mu_B(x_0)} \), but \( x_0 \notin \mathcal{B}_{\mu_B(x_0)} \). The same argument can be made to show that if \( \mu_B(x_0) > \mu_A(x_0) \) then \( \mathcal{A}_{\mu_B(x_0)} \neq \mathcal{B}_{\mu_B(x_0)} \). Ergo, the proposition is true for Case 1.

Case 2: Let \( \alpha^* > \beta^* \). If \( \mathcal{A}_{\text{max}} = \mathcal{B}_{\text{max}} \) then the proposition is true if we take \( t = \alpha^* \), since \( \mathcal{A}_t = \mathcal{A}_{\text{max}} \), but \( \mathcal{B}_t = \mathcal{B}_t = \mathcal{A}_{\text{max}} \).

Now, if we take \( \mathcal{A}_{\text{max}} \neq \mathcal{B}_{\text{max}} \) then the proposition is true if we take \( t = \alpha^* \), provided that \( \mathcal{A}_{\text{max}} \neq \mathcal{B}_t \), since \( \mathcal{B}_t = \mathcal{B}_{\text{max}} \).

Suppose that \( \mathcal{B}_t \) is constructed in a fashion such that \( \mathcal{B}_t = \mathcal{A}_{\text{max}} \). Since \( \mathcal{B}_t \supset \mathcal{B}_{\text{max}} \), we can simply choose \( t = \beta^* \). We have \( \mathcal{B}_t = \mathcal{B}_{\text{max}} \), but \( \mathcal{A}_t \supset \mathcal{A}_{\text{max}} = \mathcal{B}_t \). \( \square \)

It is a metric since \( \mathcal{H} \) is well defined by Proposition 3.1 and therefore it is easy to see that it satisfies the triangle inequality. Propositions 3.2 and 3.3 ensure identity and positive definiteness, respectively.

\( \mathcal{H} \) is a membership weighted average of the Hausdorff distances between the level sets of the two fuzzy sets. It is important to note that some level sets are modified as necessary (\( \mathcal{A}_t \)) to preserve the metric qualities.

3.2. Fuzzy Banach Space. (This section will need to go somewhere else, but I'll leave it for now)

A Banach space over a fuzzy topological vector space should behave much the same as a Banach space over a vector space. For this reason, we are able to make the following definition:

Definition 3.1. If we equip some fvs \( X \) with an \( L^p \) fuzzy norm, we define a Banach space on a fvs to be a complete normed fvs. That is, a sequence of fuzzy vectors must converge in the space. Namely, there exists an \( x \in X \) satisfying

\[
\lim_{n \to \infty} ||x_n - x||_X = 0
\]
for all sequences of fuzzy vectors \( \{x_n\} \). Moreover,
\[
\sum_{n \in \mathbb{N}} \|x_n\|_X = \sum_{n \in \mathbb{N}} x_n \text{ converges in } X
\]

4. **Fuzzy Topologies**

**Definition 4.1.** A fuzzy topology on a set \( X \) is a family \( \mathcal{F} \) of collections of fuzzy sets in \( X^n \) such that
(i) \( 0, 1 \in \mathcal{F} \)
(ii) If \( A, B \in \mathcal{F} \), \( A \cap B \in \mathcal{F} \)
(iii) If \( A_j \in \mathcal{F} \), \( \forall j \in J \), \( \bigcup_{j \in J} A_j \in \mathcal{F} \)

**Definition 4.2.** A fuzzy set \( N \) in a fts \((X, \mathcal{F})\) is called a fuzzy neighborhood of \( x \in X \) if there exists \( G \in (X, \mathcal{F}) \) such that \( G \subset N \) and \( \mu_N(x) = \mu_G(x) \).

4.1. **Fuzzy Topological Separation Axioms.** For completeness, we give new definitions of topological separation axioms for \( T_0, T_1, T_2, T_3, T_4 \).

**Definition 4.3.** A fts \((X, \mathcal{F})\) is a fuzzy \( T_0 \) space iff for any two distinct points in \((X, \mathcal{F})\), at least one of them has a neighborhood which is not a neighborhood of the other. In other words, if any two distinct points are topologically distinguishable then \((X, \mathcal{F})\) is \( T_0 \).

**Definition 4.4.** A fts \((X, \mathcal{F})\) is a fuzzy \( T_1 \) space iff every fuzzy point is a closed fuzzy set. Formally, a fts \((X, \mathcal{F})\) is a fuzzy \( T_1 \) space iff for each \( x \in X \) and each \( \lambda \in [0, 1] \) there exists \( B \in \mathcal{F} \) such that \( \mu_B(x) = 1 - \lambda, \mu_B(y) = 1 \mid y \neq x \).

**Definition 4.5.** A fts \((X, \mathcal{F})\) is a fuzzy \( T_2 \) (Hausdorff) space iff for any distinct points \( y_\lambda, x_\nu \in X \), there exists open disjoint \( U, V \in \mathcal{F} (\mu_U \cap \mu_V = \mu_{k_0}) \) with \( y_\lambda \in U \) and \( x_\nu \in V \).

**Definition 4.6.** A fts \((X, \mathcal{F})\) is fuzzy regular iff, for any point \( y_\lambda \) and any one open neighborhood \( A \), there is a fuzzy set \( B \) such that:
\[
y_\lambda \in B^0 \subset B \subset A
\]

**Definition 4.7.** A fts \((X, \mathcal{F})\) is \( T_3 \) iff it is \( T_0 \) and fuzzy regular.

**Definition 4.8.** A fts \((X, \mathcal{F})\) is a fuzzy \( T_4 \) space iff for any closed distinct fuzzy sets \( A, B \), there exists open disjoint \( U, V \in \mathcal{F} (\mu_U \cap \mu_V = \mu_{k_0}) \) with \( \mu_A < \mu_U \) and \( \mu_B < \mu_V \).

5. **Fuzzy Topological Vector Spaces**

Let \( V \) denote a vector space over the field \( \mathbb{K} \).

**Definition 5.1.** Let \( \{A_j\}_{1 \leq j \leq n} \) be a finite family of fuzzy sets in the vector space \( V \). The sum \( A = \sum_{j=1}^{n} A_j \) of the family \( \{A_j\} \) is the fuzzy set in \( V \) given membership function
\[
\mu_A(x) = \sup_{\sum_{j=1}^{n} \lambda_j \leq 0} \left( \min \{\mu_{A_j}(x)\} \right), \quad x \in V
\]
Moreover, the scalar product \( \alpha A \) of \( \alpha \in \mathbb{K} \) and \( A \in V \) is the fuzzy set with membership function
\[
\mu_{\alpha A}(x) = \begin{cases} 
\mu_A(x) : \alpha \neq 0 \\
\mu_{0 A}(x) : \alpha = 0
\end{cases}
\]
with \( \lambda = \sup_{y \in V} \mu_A(y) \).

In this way, the membership function is invariant under scaling, and thus is well-defined.

**Definition 5.2.** A fuzzy topological vector space is a vector space \( V \) over \( \mathbb{K} \), \( V \) equipped with a fuzzy topology \( \mathcal{F} \) and \( \mathbb{K} \) with the usual topology \( \mathcal{H} \), such that the two mappings
(i) \( (x, y) \to x + y \text{ mapping } (V, \mathcal{F}) \times (V, \mathcal{F}) \to (V, \mathcal{F}) \)
(ii) \( (\alpha, x) \to \alpha x \text{ mapping } (\mathbb{K}, \mathcal{H}) \times (V, \mathcal{F}) \to (V, \mathcal{F}) \)
are fuzzy continuous.

Categorically speaking, if \( \mathcal{C}(V) \) and \( \mathcal{C}(\mathcal{F}) \) are the categories of fuzzy vector spaces and fuzzy topological spaces, respectively, then the category of fuzzy topological vector spaces is such that

\[ \mathcal{C}(\mathcal{F}V) \subset \mathcal{C}(V) \cap \mathcal{C}(\mathcal{F}). \]

6. FUZZY K-DIFFERENTIATION

In order to develop the theory of Fuzzy Differential Topology, we must first define notions of fuzzy differentiation as well as fuzzy k-differentiation, using fuzzy Hausdorff Spaces. In this section, we shall give a new notion of fuzzy differentiation between fuzzy topological vector spaces, and give analogous definitions for k-differentiable fuzzy mappings between these ftvs's.

**Definition 6.1.** Let \((X, \mathcal{F})\) and \((Y, \mathcal{Y})\) be two ftvs, with a mapping \(f : X \to Y\). The mapping \(f\) is fuzzy continuous at \(x \in X\) if for every neighborhood \(V\) in \(\mathcal{Y}\) of \(y = (f(x))_\delta, 0 < \delta \leq 1\), \(f^{-1}[V]\) is a neighborhood in \(\mathcal{F}\) of \(x\), \(0 < \lambda \leq \delta\). \(f\) is said to be fuzzy open if for each open fuzzy set \(U \in \mathcal{F}\), \(f[U] \in \mathcal{Y}\).

To extend the notion of the regular derivative to functions of the form \(f : E \to F\), where \(E\) and \(F\) are fuzzy topological vector spaces, and \(A\) is a nonempty subset of \(E\). The main difficulty lies in the meaning of the quotient

\[ \frac{f(a + h) - f(a)}{h} \]

Since \(F\) is a fuzzy topological vector space, making sense of \(f(a + h) - f(a)\) is easy. Rather than attempting to define the notion of the quotient of this vector, I first turn to the notion of a directional derivative in the direction of \(u \neq 0\) \(u \in E\).

It makes sense to consider the vector \(f(a + tu) - f(a)\) where \(t \in \mathbb{R}\). The following now has a meaning:

\[ \frac{f(a + tu) - f(a)}{t} \]

The general intuition is that in \(E\), the points of the form \(a + tu\) form a line, where \(t \in \mathbb{R}\), and the image of this line defines a curve in \(F\). The curve lies on the image \(f(E)\) of \(E\) under the mapping \(f : E \to F\). The curve is defined by the mapping \(t \mapsto f(a + tu)\), from \(\mathbb{R}\) to \(F\). The directional derivative \(D_u f(a)\) defines the direction of the tangent line at \(a\) to this curve. The following definition follows

**Definition 6.2.** Let \(E\) and \(F\) be two fuzzy topological vector spaces of dimension \(n\), let \(A\) be a nonempty open subset of \(E\), and let \(f : A \to F\) be any function. For any \(a \in A\), for any \(u \neq 0\) in \(E\), the directional derivative of \(f\) with respect to the fuzzy vector \(u\), denoted by \(D_u f(a)\), is the limit, if it exists,

\[ \lim_{t \to 0, t \neq 0} \frac{f(a + tu) - f(a)}{t}, \quad U = \{t \in \mathbb{R} | a + tu \in A, t \neq 0\} \]

We say that \(f\) is differentiable at \(a\) in the direction \(u\) if the limit exists. We say that \(f\) is continuously differentiable (or \(\mathcal{C}^1\)) on \(A\) if the limit exists for all \(a \in A\) and all \(u \in E\) and if \(\partial f : (A \subseteq E) \times E \to F\) is fuzzy continuous.

**Definition 6.3.** The second derivative is the derivative of the first derivative. We differentiate \(\partial_u f(a)\) with respect to \(f\) only, in the direction \(v \in E\). The second derivative is

\[ \partial^2_{uv} f(a) = \lim_{t \to 0} \frac{\partial_u f(a + tv) - \partial_u f(a)}{t} \]

We say that \(f\) is \(\mathcal{C}^2\) if \(\partial f\) is \(\mathcal{C}^1\), which happens if and only if \(\partial^2 f\) exists and is fuzzy continuous. If \(f : A \subseteq E \to F\) we require \(\partial^2 f\) to be continuous jointly on the product as a map

\[ D^2 f : (A \subseteq E) \times E \times E \to F \]

**Definition 6.4.** The third derivative is the derivative of the second derivative. Since \(\partial^2_{uv} f(a)\) is linear separately in \(u\) and \(v\), we take only its partial derivative with respect to \(f\) in the direction \(w \in E\).

\[ \partial^3_{uvw} f(a) = \lim_{t \to 0} \frac{\partial^2_{uw} f(a + tw) - \partial^2_{uv} f(a)}{t} \]
Similarly definitions apply to higher derivatives. The $k^{th}$ derivative
\[ \partial^k_{u_1, u_2, \ldots, u_k} f(a) \]
is the map:
\[ \partial^k f : (A \subseteq E) \times E \times \cdots \times E \to F \]
defined recursively as:
\[ \partial^k_{u_1, \ldots, u_k} f(a) = \lim_{t \to 0} \frac{\partial^{k-1}_{u_1, \ldots, u_{k-1}} f(a + tu_k) - \partial^{k-1}_{u_1, \ldots, u_{k-1}} f(a)}{t} \]

We say that $f$ is of class $C^k$ if $\partial^k f$ exists and is continuous.

6.1. **Fuzzy Fréchet Derivatives.** Directional derivatives, however, present a serious problem in the fact that their definition is not sufficiently uniform. Indeed, there is no reason to believe that the directional derivatives with respect to all non-null vectors $u$ share something in common. Thus, it is possible for a function to have all directional derivatives at $a$, yet not be continuous at $a$. Two functions may have all directional derivatives in some open sets, yet their composition may not as well. Thus, we induce a more uniform notion. Assume in the below discussion that the ftvs's are normed ftv's and thus are say, fuzzy Hilbert spaces, fuzzy Banach spaces, or even fuzzy locally convex topological vector space. For this reason, all of these spaces have the notion of a cartesian product.

**Definition 6.5.** Given two ftvs's $X, Y$, We say that a linear map $f : X \to Y$ is bounded if
\[ ||f(x)||_Y \leq m||x||_X \]
with $m \in \mathbb{K}$ and we say $M = \min\{m\}$ the fuzzy operator norm of $f$.

**Definition 6.6.** Let $E$ and $F$ be two fuzzy topological vector spaces, let $A$ be a nonempty open subset of $E$, and let $f : A \to F$ be any function. For any $a \in A$, we say that $f$ is differentiable at $a \in A$ if there is a linear continuous map $L : E \to F$ and a function $\epsilon(h)$, such that
\[ f(a + h) = f(a) + L(h) + \epsilon(h)||h|| \]
for every $a + h \in A$, where $\epsilon(h)$ is defined for every $h$ such that $a + h \in A$, and
\[ \lim_{h \to 0, h \in U} \epsilon(h) = 0, \quad U = \{h \in E, |a + h \in A, h \neq 0\} \]
The linear map $L$ is denoted by $Df(a)$ or $D_{\alpha}^f$, or $df(a)$ or $f'(a)$.

**Proposition 6.1.** Let $\{(X_j, \mathcal{F}_j)\}_{j \in J}, \{(Y_j, \mathcal{F}_j)\}_{j \in J}$ be two families of ftvs's and $(X, F), (Y, \mathcal{F})$ the respective products ftv's. For each $j \in J$, let $f_j : (X_j, \mathcal{F}_j) \to (Y_j, \mathcal{F}_j)$. Then, the product mapping
\[ f = \prod_{j \in J} f_j : (x_j) \to (f(x_j)), f : (X, \mathcal{F}) \to (Y, \mathcal{F}) \]
is fuzzy continuous if each $f_j$ is fuzzy continuous for each $j \in J$

Proof. See [3] \[ \square \]

**Proposition 6.2.** Let $\{(X_j, \mathcal{F}_j)\}_{1 \leq j \leq n}, \{(Y_j, \mathcal{F}_j)\}_{1 \leq j \leq n}$ be two finite families of ftv's and $(X, \mathcal{F}), (Y, \mathcal{F})$ the respective products ftv's. For each $j, 1 \leq j \leq n$, let $f_j : (X_j, \mathcal{F}_j) \to (Y_j, \mathcal{F}_j)$. Then, the product mapping
\[ f = \prod_{j=1}^{n} f_j : (x_j) \to (f(x_j)), f : (X, \mathcal{F}) \to (Y, \mathcal{F}) \]
is fuzzy open if each $f_j$ is fuzzy open for each $1 \leq j \leq n$

Proof. See [3] \[ \square \]

**Definition 6.7.** Let $E, F$ be two ftv's and let $f$ be a mapping from $E$ to $F$ ($f : E \to F$). A continuous bounded linear mapping $\phi : E \to F, \phi \in \mathcal{L}(E, F)$ is a fuzzy derivative of $f$ at $e \in E$ provided that for all $h \in E$ it holds that:
\[ f(e + h) - f(e) = \phi(h) + o(||h||) \]

\[ ^1 \text{We use the notation } \mathcal{L}(X, Y) \text{ to denote the space of bounded linear mappings from } X \text{ to } Y. \]
Definition 6.8. In order to discuss higher order derivatives, we need to define a space of $p$-multilinear fuzzy bounded mappings, denoted $L^p(X, Y)$ with the natural isomorphism $L^p(X, Y) = L(X, L^{p-1}(X, Y))$ for $p > 1$ and $f$ of mappings
\[ F : \prod_{j=1}^{p} X_j \to Y = f : \bigotimes_{j=1}^{p} X_j \to Y \]
where each $f : X_i \to Y$ is a linear fuzzy continuous mapping.

In order to discuss higher order derivatives, we need to define a space of $YOUSUF M. SOLIMAN$

Definition 6.9. Let $E, F$ be two ftvs's, each endowed with a $T_1$ fuzzy topology. Let $f : E \to F$ be fuzzy differentiable. If $f'$ is fuzzy continuous, then we say that $f$ is of fuzzy class $C^1$. We can define maps of class $C^k$ ($k \geq 1$) inductively. The $k$th derivative $D^k f$ is defined as $D \left( D^{k-1} f \right)$ and is itself a map of $E$ into $L^k \left( E, L(E, \ldots, L(E, F)) \right)$ or simply $L^k(E, F)$. A map $f$ is said to be of class $C^k$ if its $p$th fuzzy derivative $D^p f$ exists for $1 \leq p \leq k$ and is fuzzy continuous. Notice that by definition, we may decompose a $k$-multilinear bounded fuzzy mapping into linear bounded fuzzy mappings via def 4.4:
\[ D^k f \in L^k = g : \bigotimes_{j=1}^{k} E_j \to L^k(X, Y) \]

**Proposition 6.3.** Let $E_1, E_2, F_1, F_2$ be ftvs. Suppose that $f_1 : E_1 \to F_1$ and $f_2 : E_2 \to F_2$ are fuzzy differentiable of class $C^k$. Then
\[ f_1 \times f_2 : E_1 \times E_2 \to F_1 \times F_2 \]
is fuzzy differentiable of class $C^k$.

**Proof.** This is a trivial calculation in the product topology.

7. **Fuzzy Atlas of Class $C^k$**

**Definition 7.1.** Given two ftvs's $E, F$, a continuously fuzzy differentiable map $f : E \to F$ is called a diffeomorphism if it is a bijection and its inverse $f^{-1} : F \to E$ is continuously fuzzy differentiable as well. We say that $f$ is a diffeomorphism of class $C^k$ if $Df$ is fuzzy continuous, and $D^p f$ and $f^{-p}$ exist for all $1 \leq p \leq k$.

**Proposition 7.1.** Let $E_1, E_2, V_1, V_2$ be ftvs. Suppose that $f : E_1 \to V_1$ and $g : E_1 \to V_2$ are fuzzy diffeomorphisms of class $C^k$. Then $f \times g : E_1 \times E_2 \to V_1 \times V_2$ is a fuzzy diffeomorphism of class $C^k$.

**Proof.** Obvious from propositions 5.1, 5.2, and 5.3

**Definition 7.2.** Let $X$ be a set. A fuzzy atlas $\Psi$ of fuzzy class $C^k$ on $X$ is a collection of charts $(\psi_j, A_j)$ satisfying:

1. Each $A_j$ is a fuzzy set such that $\bigcup_j A_j = \sup \mu_{A_j}(x) = 1 \quad \forall x \in X$. That is, $\mathcal{U} = \{A_j\}$ is an open fuzzy cover of $X$.
2. Each $\psi_j$ is a bijection, defined on the support of $A_j$,
   \[ \text{supp}(A_j) = \{ x \in X \mid \mu_{A_j}(x) > 0 \} \]
   which maps $A_j$ onto an open fuzzy set $\psi_j[A_j] \in E$, some ftvs. Moreover, for each index $i \in J$, $\psi_i[A_i \cap A_j]$ is an open fuzzy set in $E$.
3. The mapping $\psi_i \circ \psi_j^{-1} : \psi_j[A_j \cap A_i] \to \psi_i[A_i \cap A_i]$ is a fuzzy diffeomorphism of class $C^k$ which is well-defined since $\psi_j[A_j \cap A_i]$ and $\psi_i[A_i \cap A_i]$ are open fuzzy sets in some $E$.

Each pair $(A_j, \psi_j)$ is called a fuzzy chart of the fuzzy atlas. If a point $x \in X$ lies in the support of $A_j$, then $(A_j, \psi_j)$ is said to be a fuzzy chart at $x$.

**Proposition 7.2.** Let $\Psi$ and $\Theta$ be fuzzy $C^k$ atlases with charts $(\psi_j, U_j)$ on a set $X_1$ and $(\theta_i, V_i)$ on a set $X_2$, respectively. Then the collection of charts $(\psi_j \times \theta_i, U_j \times V_i)$ forms a fuzzy $C^k$ atlas on $X_1 \times X_2$.

**Proof.** (i) and (ii) follow directly from the proof in [23].
(iii) follows directly from proposition 5.1.
Let \((X, \mathcal{F})\) be a fuzzy topological space. Suppose there exist an open fuzzy set \(A \in X\) and fuzzy continuous bijective mapping \(\psi\), defined on the support of \(A\), mapping \(A\) into some open fuzzy set in some ftvs \(E\). We say that \((\psi, A)\) and \(\{(\psi_j, A_j)\}\) are compatible if each mapping \(\psi_j \circ \psi^{-1}: \psi[A \cap A_j] \to \psi[A \cap A_j]\) is a fuzzy diffeomorphism of class \(C^k\). Two fuzzy atlases are compatible if each chart in one is compatible with each chart in the other. In this case we write \(\Psi_1 \sim \Psi_2\) to denote an equivalence relation between the two atlases. Consequently, we may identify a unique maximal atlas \(\mathcal{F}\), which contains every atlas \(\Psi\) in some equivalence class. We call the atlas \(\mathcal{F}\) the fuzzy differential structure, and identify with it a fuzzy manifold \(M\) that is equipped with \(\mathcal{F}\). That is, the pair \((M, \mathcal{F})\) is unique.

**Definition 7.3.** Let \(\mathcal{A} \subset (\mathcal{M}, \Pi)\). We say \(\mathcal{A}\) is a fuzzy \(C^k\) submanifold is there exists \(k \in \mathbb{Z}^+\), such that each fuzzy point in \(\mathcal{A}\) belongs to the domain of a chart \((U, \psi) \in \Pi\) with

\[
U \cap \mathcal{A} = \psi^{-1}(X^{k<n})
\]

where \(X^{k<n} \subset X^n\) where the last \(n-k\) entries of a fuzzy vector are \(x_0\).

**8. Example**

Suppose \(X = S^1\) is the set of points of the unit circle in \(\mathbb{R}^2\). If \(A\) is the fuzzy set of \(S^1\) consisting of the points \((\cos t, \sin t), \ 0 \leq t < 2\pi\) with the membership function of \(A\), \(\mu_A: S^1 \to I\), defined as

\[
\mu_A(t) = 1
\]

then the function

\[
\phi_1: S^1 \to \mathbb{R}
\]

is a bijection onto an open fuzzy set of \(\mathbb{R}\) and so \((\phi_1, A)\) is a fuzzy chart for \(S^1\). If we have another fuzzy set \(B\)

\[
(cos t, \sin t), \ -\pi \leq t < \pi
\]

of \(S^1\) with the characteristic function \(\mu_B : S^1 \to I\) defined by

\[
\mu_B(t) = \frac{1}{2},
\]

then the function

\[
\phi_2 : S^1 \to \mathbb{R}
\]

is another fuzzy chart \((\phi_2, B)\) for \(S^1\).

It is easy to see that

\[
\sup \{\mu_A(x), \mu_B(x)\} = 1
\]

If we define

\[
\phi_1 = \phi_2 \text{ if } 0 \leq \phi_1 < \pi \quad \text{and} \quad \phi_2 = \phi_1 - 2\pi \text{ if } \pi \leq \phi_1 < 2\pi
\]

then the mapping \(\phi_2 \circ \phi_1^{-1}: \mathbb{R} \to \mathbb{R}\) is defined by

\[
\phi_2 \circ \phi_1^{-1}(t) = \begin{cases} 
  t & : 0 \leq t < \pi \\
  t - 2\pi & : \pi \leq t < 2\pi
\end{cases}
\]

It is clear that this mapping is a \(C^1\)-fuzzy diffeomorphism. Therefore the fuzzy charts \(A\) and \(B\) form a \(C^1\)-fuzzy atlas on \(S^1\).
9. FUZZY DIFFERENTIABLE MAPS AND FUZZY TANGENT BUNDLES

Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{C}^k$ differentiable manifolds, and $f : \mathcal{M} \to \mathcal{N}$ a map. A pair of fuzzy charts $(\psi, U)$ for $\mathcal{M}$ and $(\theta, V)$ for $\mathcal{N}$ are said to be adapted to $f$ if $f[U] \subset V$. Consequently, the map
\[
\theta f \psi^{-1} : \psi[U \cap f^{-1}[V]] \to \theta[V]
\]

is defined, and we call it the fuzzy local representation of $f$ in the given fuzzy charts, at the point $x$ if $x \in U$. We say that the map $f$ is fuzzy differentiable at $x$ if it has a local representation at $x$ which is fuzzy differentiable. Naturally, this definition makes sense since we're mapping between open fuzzy sets. Similarly, $f$ is $\mathcal{C}^k$ fuzzy differentiable at $x$ if it has a fuzzy local representation which is $\mathcal{C}^k$ fuzzy differentiable.

**Proposition 9.1.** If $f$ is $\mathcal{C}^k$, then every fuzzy local representation is $\mathcal{C}^k$.

**Proof.** Let $(\psi, U)$ and $(\theta, V)$ be fuzzy charts adapted to $f$, and suppose $f$ is $\mathcal{C}^k$. We need prove that $\theta f \psi^{-1} \in \mathcal{C}^k$. In order to do so let $y_\delta \in \psi[V]$ and set $x_\lambda = \psi^{-1}[y]$ with $0 < \lambda \leq \delta$. Let $(\psi_0, U_0)$ and $(\theta_0, V_0)$ be adapted charts to $f$ giving the $\mathcal{C}^k$ fuzzy local representation $\theta_0 f \psi_0^{-1}$ at $x$. Set $U_0$ and $V_0$ such that they are fuzzy sets with $\mu_{U_0} \leq \mu_{U}$ and $\mu_{V_0} \leq \mu_{V}$. Then,
\[
\theta f \psi^{-1} = (\theta_0 \psi_0^{-1})(\theta_0 f \psi_0^{-1})(\psi_0 \psi^{-1})
\]
in $\psi(U_0)$. Naturally $(\theta_0 \psi_0^{-1})(\psi_0 \psi^{-1}) \in \mathcal{C}^k$ so we see that the fuzzy local representation $\theta f \psi^{-1} | \psi(U_0)$ is just a composition of $\mathcal{C}^k$ maps, and is therefore $\mathcal{C}^k$ in some neighborhood of every point.

**Proposition 9.2.** Let $f : \mathcal{M} \to \mathcal{N}$ and $g : \mathcal{N} \to \mathcal{P}$ each be $\mathcal{C}^k$ maps between fuzzy $\mathcal{C}^k$ manifolds. Then $gf : \mathcal{M} \to \mathcal{P}$ is $\mathcal{C}^k$.

**Proof.** $(\psi, U)$ and $(\theta, V)$ and $(\alpha, W)$ be fuzzy $\mathcal{C}^k$ charts at $x \in X$, $f(x) \in Y$, and $g(f(x)) \in Z$, respectively. We may write the fuzzy local representation of $gf$ as
\[
ag \theta^{-1} \theta f \psi^{-1} = a(g f) \psi^{-1}
\]

By proposition 6.1, this map is $\mathcal{C}^k$ in some neighborhood of every point. It follows that $gf$ is itself $\mathcal{C}^k$.

**Corollary 9.1.** If $f$ and $g$ are fuzzy $\mathcal{C}^k$ diffeomorphisms, then $gf$ is also a fuzzy $\mathcal{C}^k$ diffeomorphism.

9.1. **Fuzzy Tangent Manifolds.** We wish to define a fuzzy tangent space analogous to the traditional definition of a tangent space. That is, given a $\mathcal{C}^{r+1}$ differentiable manifold $\mathcal{M}$, we wish for our fuzzy tangent space to be of class $\mathcal{C}^r$. We shall do so by defining our space $\mathcal{T} \mathcal{M}$ in terms of equivalence classes of appropriate fuzzy vectors. A fuzzy tangent vector to $\mathcal{M}$ (underlying fvs $E$) is an equivalence class $[x, i, y_\lambda]$ with $(x, i, y_\lambda) \in \mathcal{M} \times \Lambda \times E$ so that $x \in \mathcal{M}$ is crisp. We have the equivalence relation $[x, i, y_\lambda] \equiv [y, j, w_\lambda]$ if and only if $x = y$ and
\[
\partial (\psi_i \psi_j^{-1})(\psi_i(x))v_\lambda = w_\lambda.
\]
Therefore, $\mathcal{T} \mathcal{M}$, the fuzzy tangent space, is the set of all tangent vectors to $\mathcal{M}$. We have a map
\[
\pi : \mathcal{T} \mathcal{M} \to \mathcal{M}
\]
\[
[x, i, y_\lambda] \mapsto x
\]
that is well defined. We have $\pi^{-1}(x) = \mathcal{T}_x \mathcal{M}$, the fuzzy tangent bundle may be written as
\[
\mathcal{T} \mathcal{M} = \bigsqcup_{x \in \mathcal{M}} \mathcal{T}_x \mathcal{M}.
\]
Note that the Fuzzy tangent bundle is simply a fiber bundle with fiber $E$ and base space $\mathcal{M}$. The existence of non-trivial fiber bundles in $E$ is clear, as we may embed $\iota : S^2 \hookrightarrow E$ so that $\mathcal{TS}^2$ is non-trivial. Thus, we need make use of the locally-trivializing fuzzy neighborhoods of the fiber bundle.

For an open subset $U_j \subset \mathcal{M}$ is a $\mathcal{C}^{r+1}$ manifold if we consider the pair $(U, U|U)$. We shall show that $\mathcal{T} \mathcal{M}$ is a $\mathcal{C}^{r}$. Given a fuzzy differential structure II, we have a well-defined bijective map for any $(\psi_j, U_j) \in \Pi$
\[
\mathcal{T} \psi_j : \mathcal{T} U_j \to \psi_j(U_j) \times E \subset E \times E
\]
Thus, taking this over, \( \mathcal{M} \), \( \mathcal{T}_\mathcal{M} \) a \( \mathcal{C}^r \) manifold, since \( \partial (\mathcal{T} \psi_i) (\mathcal{T} \psi_i)^{-1} \) is well-defined by proposition 6.3, and clearly of class \( \mathcal{C}^r \). In particular, the set of charts \( \{ \psi_i, \mathcal{T} U_i \}_{i \in \Lambda} \) is a \( \mathcal{C}^r \) atlas on \( \mathcal{T} \mathcal{M} \).


**Definition 9.2.** Let \( f : \mathcal{M} \to \mathcal{N} \) be a \( \mathcal{C}^{r+1} \) map. We define a map \( \mathcal{T} f : \mathcal{T} \mathcal{M} \to \mathcal{T} \mathcal{N} \) as follows:

A local representation of the map \( \mathcal{T} f \) in the charts on \( \mathcal{T} \mathcal{M} \) and \( \mathcal{T} \mathcal{N} \) is simply the derivative of the local representation of \( f \) in the charts \( \{ \psi_i, \mathcal{U}_i \}_{i \in \Lambda} \) and \( \{ \theta_j, \mathcal{V}_j \}_{j \in \Lambda} \) on \( \mathcal{M} \) and \( \mathcal{N} \), respectively. We have the requirement that \( f (\mathcal{U}_i) \subset \mathcal{V}_j \), which may be chosen by the continuity of \( f \). Namely, if \( \mathcal{V}_j \subset f (\mathcal{U}_i) \), we may just choose open sets \( \mathcal{U}_i \cap f^{-1}(\mathcal{V}_j) \). Thus we have a \( \mathcal{C}^r \) map

\[
(\mathcal{T} f)_i : \mathcal{T} \mathcal{U}_i \to \mathcal{T} \mathcal{V}_j
\]

This is independent of \( i, j \) if we choose \( i_1, i_2 \) and consider \( x \in \mathcal{U}_i \cap \mathcal{U}_j \) and use the defined equivalence relations of tangent vectors and the chain rule (a similar trick for \( j \)). Thus, there is a well-defined map \( \mathcal{T} f : \mathcal{T} \mathcal{M} \to \mathcal{T} \mathcal{N} \), so that if \( f (x) = y \), \( \mathcal{T} f : \mathcal{T} \mathcal{M} \to \mathcal{T} \mathcal{N} \).

**Proposition 9.3.** The assignments \( \mathcal{M} \mapsto \mathcal{T} \mathcal{M} \) and \( f \mapsto \mathcal{T} f \) define a covariant functor from the category of \( \mathcal{C}^{r+1} \) to \( \mathcal{C}^r \) manifolds.

**Proof.** Clearly we have the following commutative diagram of \( \mathcal{C}^r \) and \( \mathcal{C}^{r+1} \) fuzzy manifolds:

\[
\begin{array}{ccc}
\mathcal{T} \mathcal{M} & \xrightarrow{\mathcal{T} f} & \mathcal{T} \mathcal{N} \\
\pi_\mathcal{M} \downarrow & & \downarrow \pi_\mathcal{N} \\
\mathcal{M} & \xrightarrow{f} & \mathcal{N}
\end{array}
\]

that is, \( f \circ \pi_\mathcal{M} = \pi_\mathcal{N} \circ \mathcal{T} f \).

Now suppose \( f : \mathcal{M} \to \mathcal{N} \) and \( g : \mathcal{N} \to \mathcal{K} \), then following diagram of fuzzy tangent manifolds is also commutative:

\[
\begin{array}{ccc}
\mathcal{T} (g \circ f) & \xrightarrow{\mathcal{T} (g \circ f)} & \mathcal{T} \mathcal{K} \\
\mathcal{T} f & \xrightarrow{\mathcal{T} f} & \mathcal{T} \mathcal{H} \\
\mathcal{T} g & \xrightarrow{\mathcal{T} g} & \mathcal{T} \mathcal{N}
\end{array}
\]

that is, \( \mathcal{T} (g \circ f) = \mathcal{T} g \circ \mathcal{T} f \). So that, \( \mathcal{T} 1_\mathcal{M} = 1_{\mathcal{T} \mathcal{M}} \) for the identity map \( 1_Q \) on a space \( Q \).

\( \square \)

9.3. Fuzzy Cotangent Spaces. We wish to discuss the notion of a space similar to that of a non-fuzzy cotangent space, acting on our fuzzy manifolds. Intuitively, it would make sense to have this space consist of all maps \( \eta : \mathcal{T}_x \mathcal{M} \to \mathcal{K} \) where \( \mathcal{K} \) is the underlying field of \( E \). We would also like to glue these spaces together in order to create a cotangent bundle. Namely, the fuzzy cotangent bundle will be a fiber bundle with base space \( \mathcal{M} \) and the fibers will be the dual space of the fibers of \( \mathcal{T} \mathcal{M} \), or, \( \mathcal{T}_x \mathcal{M} \). Thus, given \( x \in \mathcal{M} \), our fibers of the cotangent bundle are defined as

\[ \mathcal{T}^* \mathcal{M} = \{ \eta | \eta : \mathcal{T} \mathcal{M} \to \mathcal{K} \}. \]

Thus, taking this over, \( \mathcal{M} \),

\[ \mathcal{T}^* \mathcal{M} = \bigsqcup_{x \in \mathcal{M}} \mathcal{T}^*_x \mathcal{M}. \]
It should be noted that, just as every map \( f : \mathcal{M} \to \mathcal{N} \) induces a pushforward map \( f^* : \mathcal{T}_x \mathcal{M} \to \mathcal{T}_{f(x)} \mathcal{N} \), every pushforward map defines a pullback map

\[
f^* : \mathcal{T}_{f(x)} \mathcal{N} \to \mathcal{T}_x \mathcal{M}.
\]

Explicitly, this is defined as the relation

\[
f^*(\eta)(X_x) = \eta(f^*X_x)
\]

where \( \eta \in \mathcal{T}_{f(x)}^* \mathcal{N} \) and \( X_x \in \mathcal{T}_x \mathcal{M} \).

We may define fuzzy differential forms similar to differential forms in the standard setting of euclidean space.

**Definition 9.3.** A fuzzy differential \( p \)-form on a fuzzy differentiable manifold \( \mathcal{M} \) is a bundle section of the vector bundle of

\[
\bigwedge^p \mathcal{T}^* \mathcal{M}.
\]

Similarly, the set \( \Omega^p(\mathcal{M}) \) of all smooth \( p \)-forms, is the set of all smooth sections of

\[
\bigwedge^p \mathcal{T}^* \mathcal{M}.
\]

10. ALTERNATING FORMS ON FTVS’S

We assume that all \( A \) are \( n \)-dimensional ftvs’s containing fuzzy vectors, and whenever \( k \) appears, assume \( k \in \mathbb{N}^+ \).

Let \( S_k \) denote the permutation group of \( k \) elements with members of the group written as bijective mappings \( \sigma : \{1, \ldots, k\} \to \{1, \ldots, k\} \). Define a natural group homomorphism \( \rho : S_k \to \pm 1 \) by

\[
\rho(\sigma) = (-1)^v \quad v := \#\{(i, j) \in \{1, \ldots, k\}^2 : i > j, \sigma(i) > \sigma(j)\}
\]

**Definition 10.1.** Given a ftvs \( X = \prod_k A \), a fuzzy alternating \( k \)-form on a ftvs is a multilinear map

\[
\omega : X \to E
\]

satisfying

\[
\omega(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(k)}) = \rho(\sigma)\omega(x_1, x_2, \ldots, x_k) \quad \forall x_j \in A, \sigma \in S_k
\]

A fuzzy alternating 0-form is by definition a fuzzy number in \( E \). Denote by

\[
\Lambda^k A^* := \{ \omega : X \to E : \omega \text{ is a fuzzy altering } k\text{-form} \}
\]

for \( \omega \in \Lambda^k A^* \), \( k =: \deg(\omega) \) and is called the degree of the map \( \omega \).

**Definition 10.2.**
11. FUZZY LEBESGUE INTEGRATION

**Definition 11.1.** Let \( X \) and \( Y \) be sets, and let \( f : X \to Y \) be a bounded measurable\(^2\) mapping, and let \( A \) and \( B \) be fuzzy sets in \( X \) and \( Y \), respectively. We note that \( \mu_{f[A]}(x) : Y \to [0,1] \to [0,1] \) (see def 2.3). Let \( \Gamma := \{(x,f(x)),(f(x),\mu_{f[A]}(x)) \subseteq X \times Y \times [0,1] \} \) With respect to two partitions \( P_1 \) and \( P_2 \) of bounded measurable sets \( M_1 \subseteq X, M_2 \subseteq Y \) under a measure \( \eta \), we define

\[
\Omega_1 := \{ m \in M_1 : f_{i-1}(x) \leq f(m) < f(x) \}, \quad \Omega_2 := \{ m \in M_2 : \mu_{f[A]}(y) - 1 \leq \mu_{f[A]}(m) < \mu_{f[A]}(y) \}
\]

Now define

\[
\Psi = \begin{cases} \sum_{x \in \Omega_2} \frac{\mu_{f[A]}(x)}{\eta(M_2)} : \mu_A, \mu_B \text{ well-defined} \\ 1 : \text{else} \end{cases}
\]

\[
\int_{M_1,M_2} \Gamma \ d \eta = \lim_{||P_1, P_2|| \to \infty} \left( (\Psi) \sum_{i \in P_1} f^*(x), \eta(\Omega_1) \right)
\]

For example, if a map \( f \) does not map between fuzzy components, then \( \Psi = 1 \) and this will simply be the usual Lebesgue Integral of \( f \). Now if the membership functions are well-defined for the mapping \( f \) (that is, \( f \) is indeed mapping between two fuzzy sets) the integrand is weighted based upon the average value of \( \mu_{f[A]} \) over some domain \( M_2 \). One should note the possible computer science and physical applications of this fuzzy Lebesgue integration, for it weights a function based upon its fuzzy components.

11.1. Criteria for Integrability. We must define proper criteria to say that a fuzzy mappy is Fuzzy Lebesgue Integrable, although these are simply analogous to the usual conditions.

**Theorem 11.2.** A bounded measurable fuzzy mapping \( f \) is Fuzzy Lebesgue Integrable on bounded measurable sets \( M_1, M_2 \) if and only if, given \( \epsilon_1, \epsilon_2 > 0 \), there exist simple functions \( \tilde{f}, \tilde{g} \) and \( \check{g} \) such that

\[
f \leq \tilde{f} \leq \check{f} \quad \check{g} \leq \mu_{f[A]} \leq \tilde{g}
\]

\[
\int_{M_1} \check{f} \ d \eta - \int_{M_1} f \ d \eta < \epsilon_1
\]

and

\[
\int_{M_2} \check{g} \ d \eta - \int_{M_2} g \ d \eta < \epsilon_2
\]

**Proof.** This follows directly Definition 12.1 and Theorem X.x in [?]. \( \square \)

We shall denote fuzzy \( L^p \) space by \( L^p(\eta) \). We say that \( \Gamma \in L^p(\eta) \) if

\[
||\Gamma||_p = \left( \int_{M_1,M_2} ||\Gamma||^p \ d \eta \right)^{\frac{1}{p}} = \left( \int_{M_1} ||f||^p \ d \eta \right)^{\frac{1}{p}} \left( \int_{M_2} ||\mu_{f[A]}||^p \ d \eta \right)^{\frac{1}{p}} < \infty
\]

**Proposition 11.1.** Let \( A, B \) be fuzzy sets in some set \( X \).

If \( \Gamma_1 := \{ f, \mu_{f[A]} \} \in L^p(\eta) \) with, \( \Gamma_2 := \{ g, \mu_{g[B]} \} \in L^p(\eta) \) where \( \Gamma_1 \) has domains \( \Omega_1, \Omega_2 \), and \( \Gamma_2 \) has domains \( \Sigma_1, \Sigma_2 \), \( a \in \mathbb{R} \), then \( \Gamma_1 + \Gamma_2 \in L^p(\eta) \), and \( a \Gamma_1 \in L^p(\eta) \).

With

\[
\Gamma_1 + \Gamma_2 = \{ f + g, \mu_{f[A]} + \mu_{g[B]} \}
\]

and

\[
a \Gamma_1 = \{ af, \mu_{af[A]} \}
\]

where \( \mu_{af[A]} \) is defined in 5.1.

\(^2\)Recall \( f \) is measurable on \( A \) if \( \{ x \in A : x < r, \forall r \in \mathbb{R} \} \) is measurable with respect to a measure \( \eta \).
Proof. Let $\chi_1, \chi_2, \xi_1, \xi_2$ be the characteristic functions for $f$, $\mu_f[A]$, $g$, and $\mu_g[B]$, respectively. That is, these functions have a value of 1 if we are in the domain of the function, and 0 else.

We note that

$$
||\Gamma_1 + \Gamma_2||_p = \left( \int_{\Omega_1 \cup \Sigma_1} |\chi_1 f + \xi_1 g|^p \, d\eta \right)^{\frac{1}{p}} \left( \int_{\Omega_2 \cup \Sigma_2} |\chi_2 \mu_f[A] + \xi_2 \mu_g[B]|^p \, d\eta \right)^{\frac{1}{p}}
$$

\[ \leq 2 \left( \int_{\Omega_1 \cup \Sigma_1} |\chi_1 f + \xi_1 g|^p \, d\eta \right)^{\frac{1}{p}} \]

\[ \leq 2 \left( \int_{\Omega_1} |f|^p \, d\eta \right)^{\frac{1}{p}} + \left( \int_{\Sigma_1} |g|^p \, d\eta \right)^{\frac{1}{p}} \] (Minkowski's Inequality)

Similarly, for $\alpha \Gamma_1$, we have that

$$
||\alpha \Gamma_1||_p = \left( \int |\alpha f|^p \, d\eta \right)^{\frac{1}{p}} \left( \int |\alpha \mu_f[A]|^p \, d\eta \right)^{\frac{1}{p}}
$$

\[ = \alpha \left( \int |f|^p \, d\eta \right)^{\frac{1}{p}} \left( \int |\mu_f[A]|^p \, d\eta \right)^{\frac{1}{p}} \]

\[ \leq \alpha \left( \int |f|^p \, d\eta \right)^{\frac{1}{p}} \]

\[ < \infty \]

Proposition 11.2. If $f \in L^p(\eta)$ on some bounded measurable domain $M_1$, with $f : M_1 \to M_2$. We may define $\Gamma := \{f, \mu\}$ with

$$
\mu = \begin{cases} 1 & \text{if } y \in M_2 \\ 0 & \text{if } y \notin M_2 \end{cases}
$$

Then, $\Gamma \in L^p(\eta)$.

Proof. Notice that

$$
\left( \int_{M_1, M_2} |\Gamma|^p \, d\eta \right) = ||\mu||_p ||f||_p
$$

Since $f \in L^p(\eta)$, and $0 \leq ||\mu||_p \leq 1$ we notice that $||\mu||_p ||f||_p < \infty$. Thus, $\Gamma \in L^p(\eta)$.

11.2. A Basic Example. Let us consider a basic example in which $\Gamma := \{f, \mu\}$ where $f = x^2$

$$
\mu_A = \begin{cases} 0 & \text{if } y \leq 0 \\ y & \text{if } 0 < y < 1 \\ 1 & \text{if } y > 1 \end{cases}
$$

in the domains $M_1 := [0, 10]$, $M_2 := [0, 100]$, as depicted in the below graphs.
Naturally, \[
\int_{M_1,M_2} \Gamma \, d\eta = \left( \int_{M_1} f \, d\eta \right) \left( \int_{M_2} \mu_{f[A]} \, d\eta \right) = \left( \frac{1000}{3} \right) \left( \frac{0 + .5 + 99}{100} \right) = 331.6
\]
Part II: Fuzzy Concepts in Machine Learning
12. NOTES

12.1. locally linear embedding.
- general idea: preserve distances between points
- MDS: pairwise straightline dist, isomap: shortest-path ("geodesic") dist
- minimize reconstruction errors in lower-dimension space: \( \min \epsilon(W) = \sum \|X_i - \sum W_{ij}X_j\|^2 = \min \) sum squared distances between datapoints and reconstructions
- process: (1) assign k nearest neighbors (2) represent every point in terms of weighted neighbors \( W_{ij} \) (3) compute low-dimensional embedding which best preserves \( W \)
- issue: why is this right. yes it preserves local geometry, but why should neighborhood-preserving tell you anything about the global structure? also, knn is not a confident representation of open sets in \( M \), in fact may not even provide a good cover
- idea: show that these lower-dim reps are \( C^k \)-compatible
- this is only as strong as knn. to pick which neighbors to include is equivalent to having a great clustering in the first place, all this does is nonlinearly warp the space to make clusterings more obvious to e.g. SVM.
- also, similarity to self-organizing maps? proof sketch: this is more powerful b/c local organization can budge global organization away from optimum? but, objective value might depend on order in which reps are optimized, i.e. how do you co-optimize reps which may be on the boundary of multiple open sets.
- probably, noise LLE is best way to go.

12.2. MDS.
- equivalent to PCA when distances are euclidean
- todo

12.3. Isomap.
- MDS with knn determined by shortest geodesic path.
- speedups using dijkstra, fibonacci heaps (sleator + tarjan)

12.4. noisy ideas.
- weight matrix in LLE is influenced by membership function. we include points in a cluster to degree of membership
- naive: co-optimize in \( u(x) \times y \) for best lower-dimensional representation. issue: we could just throw out all points, becoming zero vectors.
- naive solution: modify objective function to penalize information loss... but how? idea: fit GMM and minimize peak loss. but extremely sketchy and unbelievable. also, \( O(k^2n + kmn) \times O(\text{whatever}) \) runtime is bad.
- better idea: minimize KL-divergence loss. i.e. points with high divergence are penalize the objective more if lost. but: what heuristics to compute KL divergence. i.e. divergence from neighbors, total covariance, etc.
- non-naive: optimize in \( u(x) \times y \), but assign objective \( y \) as usual to goodness of representation, while objective \( u(x) \) is clustering goodness minus information loss penalty
- better idea: interpret LLE as k-PCA with data-dependent kernel. optimizing embedding map for best separation is well-defined, and moreover we can make statistical statement by covariance measure. objective gradient will involve gradient for \( M \), which is where the differential structure comes in.
- TODO

13. INTRODUCTION: APPLICATIONS OF FUZZY DIFFERENTIABILITY TO NOISY MANIFOLD LEARNING

A classical problem in data analysis is that of dimensionality reduction. The problem is loosely phrased as follows. Assume some observation set \( \mathcal{O} \) lies on an \( m \)-dimensional manifold \( \mathcal{M} \) embedded in \( \mathbb{R}^n \) where \( n > m \) (that is, assume WLOG that every open set \( U \subset \mathcal{M} \) is homeomorphic to \( \mathbb{R}^m \); practically speaking, these open sets may be different-dimensional, but this is addressed later). We wish to construct a deformation from \( \mathcal{M} \) to
the $\mathbb{R}^d$ plane, where $d \leq m$, in a manner that preserves neighborhood structure in some sense. This condition is somewhat stronger than bicontinuity. It’s insufficient to construct a one-to-one assignment; we wish to preserve the local geometry as well.

The motivation, twofold, is as follows. Firstly, planar data is in some sense easier to visualize, though with the introduction of very high dimensional data, this ceases to be the chief application. Second, the more interesting application is as a preprocessing stage in a clustering pipeline. Under the assumption that the data is parameterized by some (lower-dimensional) vector of hidden variables – that presumably induces a natural clustering of these hidden variables, composed with linear coupling with background noise. The stronger assumption is that this nonlinear coupling determines the intrinsic curvature of the manifold. We motivate this assumption with more rigor in the following sections. In this perspective, PCA flattens and warps manifolds which are the result of linear couplings of hidden variables, transforming the sheared hidden space into one more closely aligned with a planar lattice.

For observations which are linear combinations of hidden variables, PCA is optimal. The issue arises when we remove this assumption. With allowed arbitrary coupling functions, the problem, which may be interpreted as a manifestation of the "no free lunch" theorem, is that the conditions for determining point membership in specific neighborhoods – that is, the topology – are unclear. For example, Locally Linear Embedding assumes an $\epsilon$-ball euclidean metric topology; Isomap does the same with a geodesic metric; and more recently, methods of persistent homology study invariants in the Betti numbers of a parameterized $\epsilon$-ball metric topology.

We study the first phase of the problem – adjusting predicted neighborhood structure in the presence of independently distributed noise. This issue is not equivalent to the no free lunch problem, in that an independent noise source can be strongly reasoned about via known statistical measures and criteria, a la CLT. However, the problem of optimizing neighborhood reconstruction with the added parameter of point membership in the observation set has traditionally been ill-defined; with the formal introduction of fuzzy differentiation and integration from Part I, however, we may proceed.

14. **Naive Noise-resistant Locally Linear Embedding**

Consider the canonical formulation of locally linear embedding; that is, given the metric topology induced by $\epsilon$, reconstruct every point in $\mathcal{M}$ as a linear combination of its neighbors. This is a weight matrix $W$ such that $p = \sum_j W U / p$, where $p$ is in an open set $U \subset \mathcal{M}$ and $j$ indexes columns. Suppose we construct a lower-dimensional embedding of $\mathcal{M}$ via a transformation $T$; the reconstruction problem is

$$\min_{W} \sum_i \left\| X_i - \sum_j W \circ TX_j \right\|^2$$

for all $x_j \neq x_i \in U$.

Now consider, in the embedded space, the additional constraint where points have membership determined by a PDF $P(x, \theta)$ representing background noise, as well as the projection $\pi_\theta : P(x, \theta) \rightarrow \mu(x)$ which we’ll make use of in the following sections. This first variant we consider purely as an exercise in thinking about noisy manifold embeddings, though the construction is nearly identical to the non-naive case, and provides useful intuition.

14.1. **Formalization.** We denote this case as naive since we assume $P$ shifts the means uniformly; that is, $f(\mathcal{O} - \mu \circ \mathcal{O})$ is approximated by a normal distribution for some positive semidefinite error function $f : V \rightarrow \mathbb{R}^n$. We begin by constructing a pairwise potential function $\mathcal{U}$ ...

15. **Computational Foundations**

Denote $\mathcal{O}$ the observation point-cloud, embedded in a vector space $V$ with the standard basis, with $\dim V = n$. As is usual in literature we proceed form an (assumed unique) triangulation of the cloud, denote as $\mathcal{T}$ (polynomial; cf. corrected Voronoi). For now, we let this simplicial complex be an open cover of $\mathcal{M}$. This may seem an obvious construction, but it is tangibly motivated by two reasons. Firstly, no information about the space has been lost; the simplicial complex enforces nearest-neighbor adjacency, but retains the dimension of the tangent spaces at all points (though this not yet a smooth structure), letting us fully reconstruct the space to original precision. Second, in some sense this is the finest topology we can induce on $\mathcal{O}$ in the original space, if we use the usual
metric on \(V\). Any future topology that we consider for the sake of dimensionality reduction will also be metric (or pseudometric (i.e. using the Kullback-Leibler pseudometric)) and so all future constructions will build upon this complex.

**TODO...**

16. **STRONGER REASONING VIA INFORMATION GEOMETRY**

We now consider the problem under the relaxed assumption: we require only that \(P(x, \theta)\) is continuous and infinitely differentiable. Clearly it is no longer possible to exploit the Gaussian factorization used in Naive NLLE; instead we now turn to geometric information measures to construct inclusion sets \(U'\) for open sets of \(\Theta\) with reasonable confidence. To do so, we begin by inducing a metric topology \(\mathcal{T}\) over the simplicial complex. The measure we use is a Taylor approximation of the Kullback-Leibler distance (a pseudomeasure of distance between two probability distributions, interpreted as change in entropy)a

**TODO**

16.1. **Criteria for confident membership estimation.** TODO

16.2. **Various interesting isometries.** TODO

17. **MANIFOLDS OF NON-UNIFORM DIMENSIONALITY**

**TODO**

18. **CONFIDENT NOISE-RESISTANT EMBEDDINGS**

**TODO**

19. **ALGORITHM AND RELEVANT HEURISTICS**

20. **ANALYSIS**

20.1. **Convergence and Runtime.**

20.2. **Addressing assumptions.**

21. **APPROXIMATIONS FOR DISCONTINUOUS MEMBERSHIP AND OTHER PATHOLOGIES**

**REFERENCES**

[2] Marrio Ferraro, David Foster, \(C^1\) *fuzzy manifolds*, Fuzzy sets and systems, 1992

E-mail address: ysoliman@cmu.edu