Evolutions of Logifold Structures on Measure Spaces

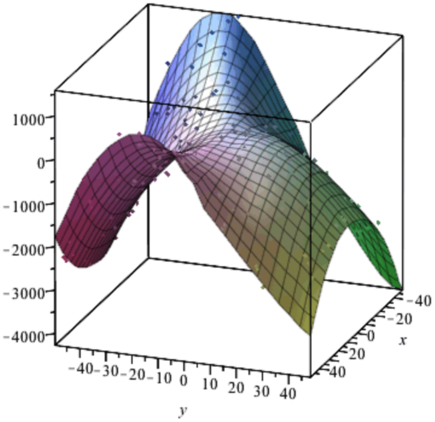
Inkee Jung
Boston University

May 28 - 31

Based on Logifold structure on measure space
joint work with Siu-Cheong Lau (arXiv:2405.05492)
BU-Keio-Tsinghua Workshop 2024 on
Differential Equation Dynamical System and Applied Mathematics

“Manifold” in Data Science

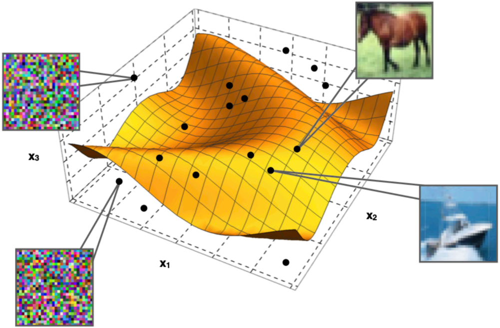
Interpolation



(Image from Maple16)

"Manifold" in Data Science

High-dimensional analogue of 2 dimensional surface in $\mathbb{R}^N$



(Image from Sebastian Goldt, Marc Mézard, Florent Krzakala, and Lenka Zdeborová)

Manifold : Local-to-Global Principle

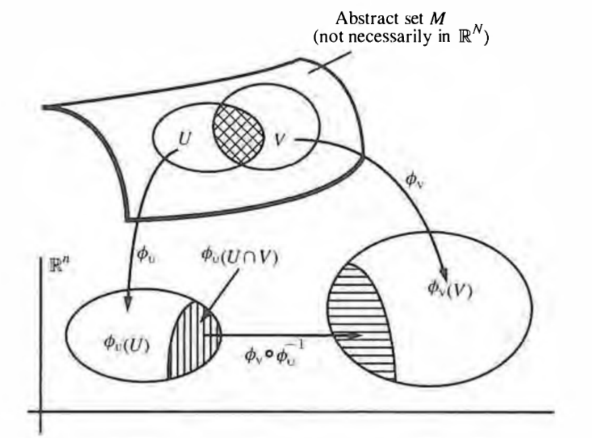
Locally Euclidean Space $(M, \mathcal{U})$ with collection of local data $\mathcal{U} = \{(U_\alpha, \Phi_\alpha)\}$

- Modeling Spacetime by Einstein's theory of relativity
- Local-to-Global principle

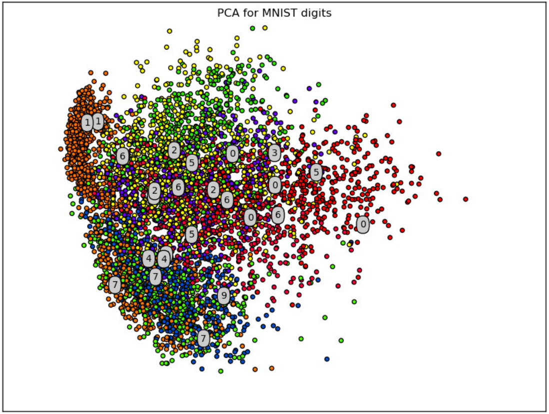
Manifold : Local-to-Global Principle

Locally Euclidean Space $(M, \mathcal{U})$ with collection of local data $\mathcal{U} = \{(U_\alpha, \Phi_\alpha)\}$

- Modeling Spacetime by Einstein's theory of relativity
- Local-to-Global principle

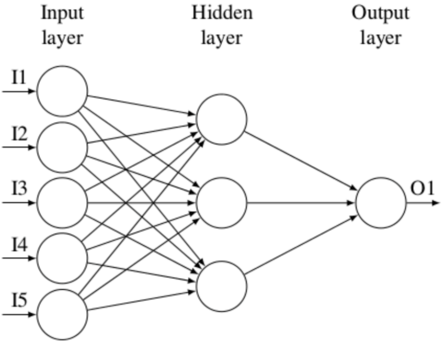


Is it really best formulation for dataset?



PCA for MNIST with 2 most important dimensions.
(created by Vijayasaradhi Indurthi)

Neural Network in Data Science



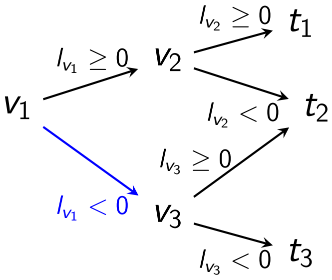
$$f = \sigma_2 \circ L_2 \circ \sigma_1 \circ L_1$$

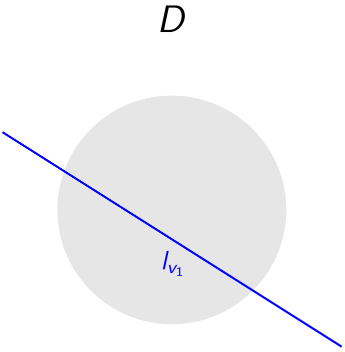
- Network models gain tremendous success in describing datasets
- In classification task, last layer is index-max function.

Linear Logical Function : Local chart

Motivated from Neural Network.

Example: Directed graph G & Set of affine maps $L = \{l_{v_1}, l_{v_2}, l_{v_3}\}$

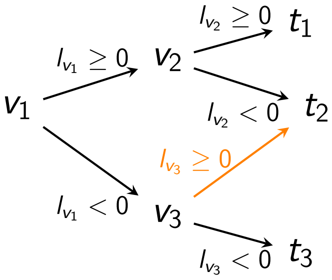


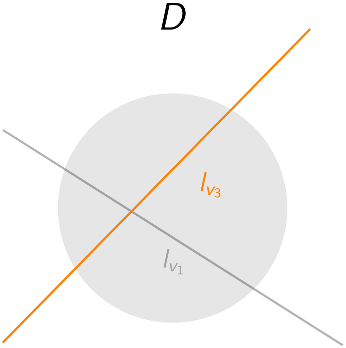


Linear Logical Function : Local chart

Motivated from Neural Network.

Example: Directed graph G & Set of affine maps $L = \{l_{v_1}, l_{v_2}, l_{v_3}\}$

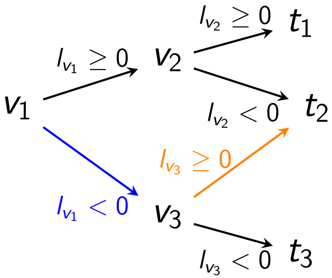


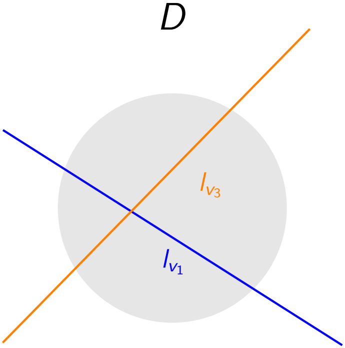


Linear Logical Function : Local chart

Motivated from Neural Network.

Example: Directed graph G & Set of affine maps $L = \{l_{v_1}, l_{v_2}, l_{v_3}\}$

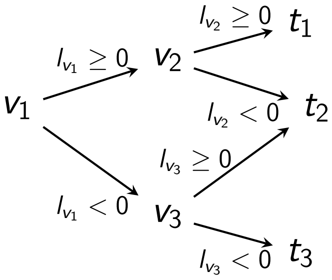


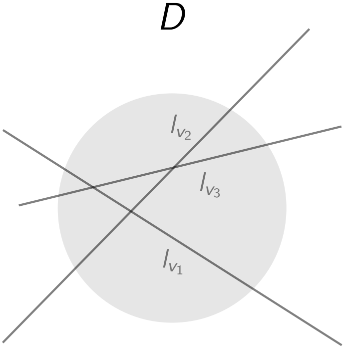


Linear Logical Function : Local chart

Motivated from Neural Network.

Example: Directed graph G & Set of affine maps $L = \{l_{v_1}, l_{v_2}, l_{v_3}\}$





Linear Logical Function : Local chart

- Measurable set $D \subset \mathbb{R}^n$, Finite set T.
- Directed finite graph G without cycle
- Affine maps

$$L = \{l_v : v \text{ is a vertex with more than one outgoing arrows}\}$$

Definition

$f_{G,L} : D \to T$ is a linear logical function of (G, L) if $l_v \in L$ are affine linear functions whose chambers in D are one-to-one corresponding to the outgoing arrows of v.
(G, L) is called a linear logical graph.

Linear Logical Function : Local chart

- Measurable set $D \subset \mathbb{R}^n$, Finite set T.
- Directed finite graph G without cycle
- Affine maps

$$L = \{l_v : v \text{ is a vertex with more than one outgoing arrows}\}$$

Definition

$f_{G,L} : D \to T$ is a linear logical function of (G, L) if $l_v \in L$ are affine linear functions whose chambers in D are one-to-one corresponding to the outgoing arrows of v.
(G, L) is called a linear logical graph.

Linear logical function as "local chart" of dataset.

Linear logical function in mathematical logic : Definable function.

Definition (Lou van den Dries)

A structure S on the real line consists of a boolean algebra S_n of subsets of $\mathbb{R}^n$ for each $n = 0, 1, \ldots,$ such that

- $\{x \in \mathbb{R}^n : x_i = x_j\}, 1 \leq i < j \leq n \in S_n$.
- Closed under Cartesian product.
- Closed under projection ($A \in S_{n+1} \to \pi(A) \in S_n$).
- $\{(x, y) \in \mathbb{R}^2 : x < y\} \in S_2$.

For instance:

$$\Phi := \{(x, y) \in X \times Y : \phi(x, y)\}$$
$$\pi_Y(\Phi) = \{y \in Y : \exists x \phi(x, y)\}$$

Linear logical function in mathematical logic : Definable function.

Definition (Lou van den Dries)

A structure S on the real line consists of a boolean algebra S_n of subsets of $\mathbb{R}^n$ for each $n = 0, 1, \ldots,$ such that

- $\{x \in \mathbb{R}^n : x_i = x_j\}, 1 \leq i < j \leq n \in S_n$.
- Closed under Cartesian product.
- Closed under projection $(A \in S_{n+1} \to \pi(A) \in S_n)$.
- $\{(x, y) \in \mathbb{R}^2 : x < y\} \in S_2$.

- $Def(\mathcal{A})$: The smallest structure on the real line contatining a collection $\mathcal{A}$
- $f : A \to B$ is definable if its graph is definable.

Linear logical function in mathematical logic : Definable function.

- A structure S : Geometric description of “definable sets” in mathematical logic.

Linear logical function in mathematical logic : Definable function.

- A structure S : Geometric description of “definable sets” in mathematical logic.
- A structure S is o-minimal if every *definable* subset is finite unions of intervals and points.

Linear logical function in mathematical logic : Definable function.

- A structure S : Geometric description of "definable sets" in mathematical logic.
- A structure S is *o*-minimal if every *definable* subset is finite unions of intervals and points.
- A *semilinear* set of $\mathbb{R}^n$: Finite unions of

$$\{x \in \mathbb{R}^n : f_1(x) = \cdots = f_k(x), g_1(x) > 0, \ldots, g_l(x) > 0\}$$

with affines f_i and g_j.

Linear logical function in mathematical logic : Definable function.

- A structure S : Geometric description of "definable sets" in mathematical logic.
- A structure S is o-minimal if every *definable* subset is finite unions of intervals and points.
- A *semilinear* set of $\mathbb{R}^n$: Finite unions of

$$\{x \in \mathbb{R}^n : f_1(x) = \cdots = f_k(x), g_1(x) > 0, \ldots, g_l(x) > 0\}$$

  with affines f_i and g_j.
- Semilinear set forms o-minimal structure.

Linear logical function in mathematical logic : Definable function.

- A structure S : Geometric description of "definable sets" in mathematical logic.
- A structure S is o-minimal if every *definable* subset is finite unions of intervals and points.
- A *semilinear* set of $\mathbb{R}^n$: Finite unions of

$$\{x \in \mathbb{R}^n : f_1(x) = \cdots = f_k(x), g_1(x) > 0, \ldots, g_l(x) > 0\}$$

with affines f_i and g_j.
- Semilinear set forms o-minimal structure.

Theorem

A function $f : D \to T$ for a finite set T where $D \subset \mathbb{R}^n$ is semilinear if and only if it is a linear logical function.

Proof of idea : Systematically construct one-to-one correspondence between the two categories of functions.

Universality of Linear logical function

- $D \subset \mathbb{R}^N$ with $\mu(D) < \infty$, where μ is the Lebesgue measure.
- T is finite

Theorem

For any (Lebesgue) measurable function $f : D \to T$, we have a linear logical function that approximates to f.

Universality of Linear logical function

- $D \subset \mathbb{R}^N$ with $\mu(D) < \infty$, where μ is the Lebesgue measure.
- T is finite

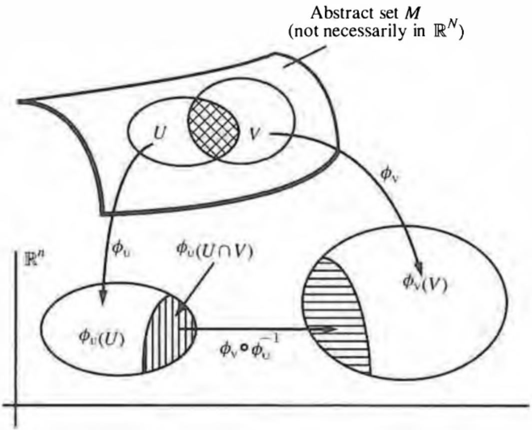
Theorem

For any (Lebesgue) measurable function $f : D \to T$, we have a linear logical function that approximates to f.

Corollary

There exists a family $\mathcal{L}$ of linear logical functions $L_i : D_i \to T$, where $D_i \subset D$ and $L_i \equiv f|_{D_i}$, such that $D \setminus \bigcup_i D_i$ is measure zero set.

Logifold


Abstract set M
(not necessarily in $\mathbb{R}^N$)
U
V
ϕ_V
ϕ_U
$\mathbb{R}^n$
$\phi_U(U \cap V)$
$\phi_V(V)$
$\phi_U(U)$
$\phi_V \circ \phi_U^{-1}$

Logifold

Definition

A linear logifold is a pair $(X, \mathcal{U})$ such that

- X is a set equipped with a σ-algebra and a corresponding measure μ.
- $\mathcal{U}$ is a collection of pairs (U_i, ϕ_i).
- U_i are subsets of X such that $\mu(U_i) > 0$ and $\bigcup_i U_i = X$.
- ϕ_i are isomorphisms (of measure spaces) between U_i and the graphs of linear logical functions $f_i : D_i \to T_i$.
- $D_i \subset \mathbb{R}^{n_i}$ are a measurable subsets and T_i are finite sets.

Example of logifolds

$\mathcal{M} \subset \mathbb{R} \times \{0, 1\}$

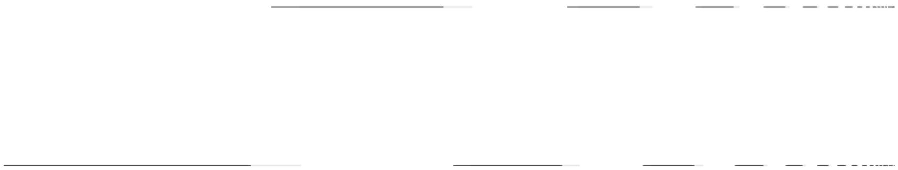
Example of logifolds

$\mathcal{M} \subset \mathbb{R} \times \{0, 1\}$

Machine learning on $\{\text{Logifolds}\}$

Example of logifolds

$\mathcal{M} \subset \mathbb{R} \times \{0, 1\}$

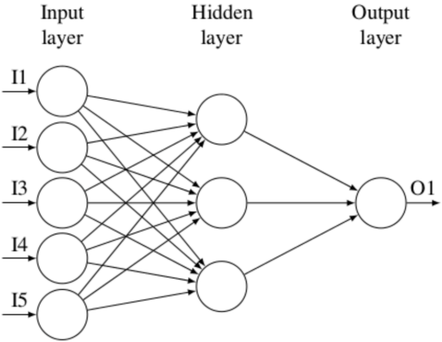


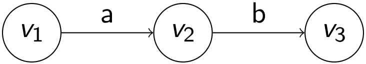
Machine learning on $\{\text{Logifolds}\}$

$? \longrightarrow \{\text{Logifolds}\}$

Motivation : Framed Quiver Moduli Space

- Q : Directed graph
- $E(Q)$: Arrows of Q
- $V(Q)$: Vertices of Q





A_3 Quiver

Motivation : Framed Quiver Moduli Space

Theorem (Nakajima('96), Reineke('08))

Framed quiver moduli space is smooth compact variety.

Motivation : Framed Quiver Moduli Space

Theorem (Nakajima('96), Reineke('08))

Framed quiver moduli space is smooth compact variety.

Theorem (Lau & Jefferys'21))

Framed quvier moduli space has the natural kähler metric.

Motivation : Framed Quiver Moduli Space

Theorem (Nakajima('96), Reineke('08))

Framed quiver moduli space is smooth compact variety.

Theorem (Lau & Jefferys'21))

Framed quvier moduli space has the natural kähler metric.

$$\text{Framed } A_3 \text{ moduli space} \longrightarrow \{\text{Logifolds}\}.$$

Motivation : Framed Quiver Moduli Space

Theorem (Nakajima('96), Reineke('08))

Framed quiver moduli space is smooth compact variety.

Theorem (Lau & Jefferys'21))

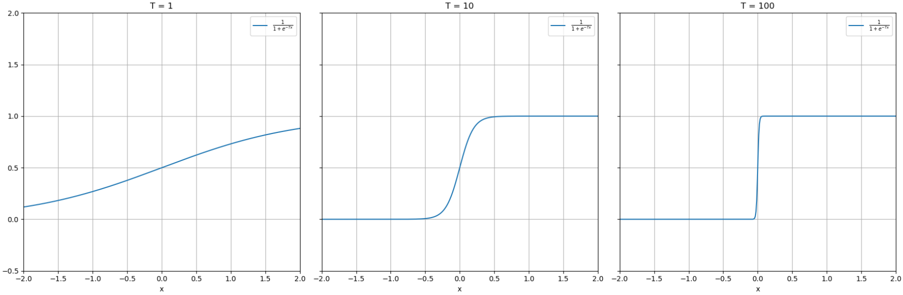
Framed quvier moduli space has the natural kähler metric.

$$\text{Framed } A_3 \text{ moduli space} \longrightarrow \{\text{Logifolds}\}.$$

Problem : Trivial gradient

Motivation : Non-archimedean analysis

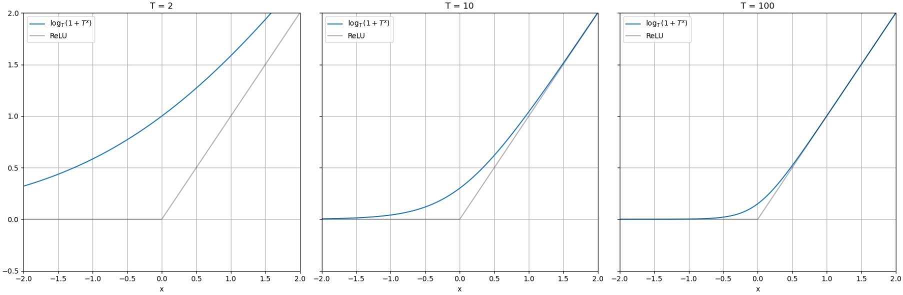
Introduce formal parameter T to logistic functions.



$$\lim_{T\to\infty} \frac{1}{1+e^{-Tx}} = \begin{cases} 1 & (x>0) \\ 0 & (x<0) \end{cases}$$

Motivation : Non-archimedean analysis

Introduce formal parameter T to logistic functions.



$$\lim_{T\to\infty} \log_T(1+T^x) = \begin{cases} x & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

Motivation : Non-archimedean analysis

Introduce formal parameter T to logistic functions.
For $T > 1$,

- $\sigma_T(x) := \frac{1}{1+e^{-Tx}}$.
- $\mathrm{ReLu}_T(x) := \log_T(1 + T^x)$.
- $\mathrm{SoftMax}_T(x_1, \ldots, x_n) := \left(\frac{T^{x_j}}{\sum_i T^{x_i}} : j = 1, \ldots, n \right)$
- Smooth, non-trivial gradient in each T

$$\frac{\partial \sigma_T(x)}{\partial x} = \frac{Te^{Tx}}{\left(e^{Tx} + 1\right)^2}$$

$$\frac{\partial \mathrm{ReLu}_T(x)}{\partial x} = \sigma_T(x)$$

$$D\left(\mathrm{SoftMax}_T\right) = \left[\log T \cdot \left(\mathrm{SoftMax}_T\right)_j \cdot \left(\delta_{ij} - T^{x_i}\right)\right]_{i,j}$$

Fuzzy linear logical function and fuzzy linear logifold

Definition

A fuzzy linear logifold is a tuple $(X, \mathcal{P}, \mathcal{U})$, where $(X, \mathcal{U})$ be a logifold and

- $\mathcal{U}$ is a collection of tuples (ρ_i, ϕ_i, f_i)
- $\rho_i : X \to [0, 1]$ describe fuzzy subsets of X with $\sum_i \rho_i \leq 1_X$
- $U_i = \{x \in X : \rho_i(x) > 0\}$ be the support of ρ_i

Fuzzy linear logical function and fuzzy linear logifold

Definition

A fuzzy linear logifold is a tuple $(X, \mathcal{P}, \mathcal{U})$, where $(X, \mathcal{U})$ be a logifold and

- $\mathcal{U}$ is a collection of tuples (ρ_i, ϕ_i, f_i)
- $\rho_i : X \to [0, 1]$ describe fuzzy subsets of X with $\sum_i \rho_i \leq 1_X$
- $U_i = \{x \in X : \rho_i(x) > 0\}$ be the support of ρ_i

In classification problems,

- $X = \mathbb{R}^n \times T$
- $\mathcal{P} : X \to [0, 1]$ describes how likely an element of $\mathbb{R}^n \times T$ is classified as 'yes'
- ρ_i be the 'weight' (or 'certainty') of the corresponding linear logical interpretation.

Experimental Results : Evolution of fuzzy linear logifolds

Dataset : CIFAR10

Model structures : Version 1 and 2 of ResNet20 and ResNet56

Certainty Threshold	Accuracy	Coverage
0	0.8316 ± 0.0160	1
0.9526	$\mathbf{0.8316 \pm 0.0160}$	$\mathbf{0.6728 \pm 0.0507}$
0.9975	0.8316 ± 0.0160	0.1130 ± 0.0484

Table: Logifold with single chart.

Certainty Threshold	Accuracy	Coverage
0	0.9304	1
0.9526	**0.9334**	**1**
0.9975	0.9290	0.6965

Table: Evolved Logifold.