Functorial Manifold Learning

Dan Shiebler

University of Oxford

July 14, 2021

Manifold Learning

- Suppose we have a finite set of points X sampled from some larger space $\mathbf{X}$ according to some probability measure $\mu_{\mathbf{X}}$ over $\mathbf{X}$.
- **Manifold Learning** techniques construct $\mathbb{R}^m$-embeddings for the points in X, which we interpret as coordinates for the support of $\mu_{\mathbf{X}}$.

Why Functoriality?

- Directly model which invariances algorithms preserve.
- Expose similarities and hierarchies between algorithms based on the categories they are functorial over.
- Derive new algorithms that satisfy functorial constraints.

Background: Finite Pseudo-Metric Spaces

- We represent datasets as finite pseudometric spaces (X, d_X)
- The morphisms in $\mathbf{PMet}$ are **non-expansive functions** $f : (X, d_X) \to (Y, d_Y)$ where $d_Y(f(x_1), f(x_2)) \leq d_X(x_1, x_2)$.

Background: Clustering Functor

- The objects in the category $\mathbf{Cov}$ are sets $(X, \mathcal{C}_X)$ where $\mathcal{C}_X$ is a (non-nested flag) cover of X
- Morphisms are **refinement preserving functions** $f : (X, \mathcal{C}_X) \to (Y, \mathcal{C}_Y)$ such that if S in $\mathcal{C}_X$ there exists S' in $\mathcal{C}_Y$ such that $f(S) \subseteq S'$.
- A **clustering functor** $\mathbf{PMet} \to \mathbf{Cov}$ maps a finite pseudometric space (X, d_X) to $(X, \mathcal{C}_X)$.

Background: Hierarchical Clustering Functor

- Sometimes it is useful to group data at different scales.
- A hierarchical clustering functor $\mathbf{PMet} \to \mathbf{Cov}^{(0,1]^{op}}$ maps (X, d_X) to the map $(0, 1] \to (X, \mathcal{C}_X)$.

Manifold Learning

- A **pairwise embedding optimization problem** is a tuple $(n, m, \{l_{ij}\})$ that represents:
  - **Find** an $n \times m$ real valued matrix A in $\mathbb{R}^{n*m}$ that minimizes $\sum_{\substack{i \in 1 \dots n \\ j \in 1 \dots n}} l_{ij}(\|A_i - A_j\|)$.
- **Manifold learning algorithms** map pseudometric spaces (X, d_X) to pairwise embedding optimization problems $(n, m, \{l_{ij}\})$.

Example: Metric Multidimensional Scaling

Metric Multidimensional Scaling maps (X, d_X) to:

Find $A \in \mathbb{R}^{n*m}$ that minimizes $\sum_{\substack{i \in 1 \ldots n \\ j \in 1 \ldots n}} (d_X(x_i, x_j) - \|A_i - A_j\|)^2$

This is a search problem in $\mathbb{R}^{n*m}$ that is parameterized by (X, d_X).

Category $\mathbf{FLoss}$ of Manifold Learning Optimization Problems

- In the preorder category $\mathbf{FLoss}$ objects are pairwise embedding optimization problems
- $\mathbf{FLoss}$ is a multiscale extension of the simpler preorder $\mathbf{L}$ in which $(n, m, \{l_{ij}\}) \leq (n', m, \{l'_{ij}\})$ iff for any $x \in \mathbb{R}_{\geq 0}, i, j \in \mathbb{N}$ we have $l_{ij}(x) \leq l'_{ij}(x)$

Manifold Learning Taxonomy

We can classify manifold learning algorithms by the subcategory of $\mathbf{PMet}$ over which they are functorial:

- $\mathbf{PMet}_{isom}$: Invariant to isometries (e.g. UMAP)
- $\mathbf{PMet}_{bij}$: Invariant to bijections (e.g. k-Vertex-Connected Scaling)
- $\mathbf{PMet}_{sur}$: Invariant to surjections (e.g. IsoMap, MMDS)
- No functors out of $\mathbf{PMet}$

Manifold Learning Factors Through Clustering

- Constructing a pairwise embedding optimization problem from a pseudometric space requires forming groups of points at different strengths.
- Every manifold learning functor $\mathbf{PMet} \xrightarrow{M} \mathbf{FLoss}$ factorizes through some hierarchical clustering functor $\mathbf{PMet} \xrightarrow{H} \mathbf{Cov}^{(0,1]^{op}} \xrightarrow{L} \mathbf{FLoss}$

Example: Metric Multidimensional Scaling

We can define a functor $MDS : \mathbf{Cov}^{(0,1]^{op}} \to \mathbf{FLoss}$ such that the composition $MDS \circ \mathcal{ML}$ maps (X, d_X) to the embedding optimization problem where $l_{ij}(a) = (d_X(x_i, x_j) - a)^2$.

Universality of Manifold Learning Algorithms

- The δ-Vietoris-Rips Complex of (X, d_X) contains the simplex $x_1, x_2, ..., x_n$ if $d(x_i, x_{i+1}) \leq \delta$
- $\mathcal{SL}$ clustering functor maps (X, d_X) to the connected components of its δ-Vietoris-Rips Complex
- $\mathcal{ML}$ clustering functor maps (X, d_X) to the maximal faces of its δ-Vietoris-Rips Complex

For any non-trivial manifold learning functor $M = L \circ H$ we have:

$$L \circ \mathcal{ML} \leq M \leq L \circ \mathcal{SL}$$

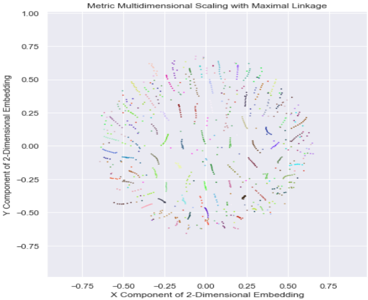
Manifold Learning Stability Properties

- Suppose we have a manifold learning functor M and the ϵ-isometric pseudo-metric spaces $(X, d_X), (Y, d_Y)$
- Suppose A_X, A_Y respectively minimize the loss functions $\mathbf{l}_{M(X,d_X)}$ and $\mathbf{l}_{M(Y,d_Y)}$
- We can use functoriality to bound $\mathbf{l}_{M(X,d_X)}(A_Y)$ in terms of $\mathbf{l}_{M(X,d_X)}(A_X)$

Functorial Recombination

- The functorial perspective on manifold learning provides a natural way to produce new manifold learning algorithms by recombining the components of existing algorithms.
- For example, we can swap $\mathcal{ML}$ with $\mathcal{SL}$ in the Metric Multidimensional Scaling functor $MDS \circ \mathcal{ML}$ to form the new Single Linkage Scaling functor $MDS \circ \mathcal{SL}$ that encourages chained embeddings.

Functorial Recombination



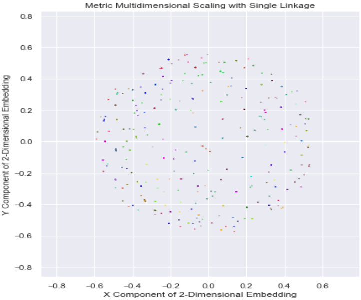


Figure: Each color indicates a unique DNA sequence at different mutation stages. Note that the $MDS \circ \mathcal{SL}$ objective (on the right) embeds sequences in the same mutation list more closely together than $MDS \circ \mathcal{ML}$ (on the left).

Thank You