Maps:

Useful in multi-variate data analysis

\( f_1: \mathbb{X} \rightarrow \mathbb{R} \) : \( f_1 \) is scalar valued function

\( f_2: \mathbb{X} \rightarrow \mathbb{R}^d \) : \( f_2 \) is vector valued function.

Alternatively, we can construct a vector of multiple scalar functions \( f = (f_1, f_2, \ldots, f_d) \).

If \( \mathbb{X} \) is a point cloud or subset of \( \mathbb{R}^d \) (\( d > 1 \)) then the domain is High Dimensional.

We'll make the distinction where High-dimensional \( \Rightarrow \) domain multi-variate \( \Rightarrow \) range for a function.

Let \( \mathbb{X} \): high-dim. point cloud with some metric/distance \( d \) \( x, y \in \mathbb{X}, \quad d(x, y) \) is computed. For this data set we can perform

\[ \rightarrow \text{Clustering: } \text{K-means, K-median, } \text{Hierarchical/graph-based} \]

\[ \rightarrow \text{Dimensionality Reduction: } \text{PCA, Factor analysis, Eigenmap, LLE, JLT etc.} \]

When we have \( y = f(x) \) \( f: \mathbb{X} \rightarrow \mathbb{R}^d \) \( d \geq 1 \)

we can perform regression \( y \)

\[ \rightarrow \text{linear regression} \]

\[ \rightarrow \text{Logit} \]

\[ \rightarrow \text{multivariate regression} \]

\[ \rightarrow \text{mapper: discretization of Reeb graph/Reeb space, "Clustering"} \]

\[ \rightarrow \text{easily integrated into machine learning.} \]
**X**: topological space \(\text{ (point cloud, manifold) }\)

**U**: open cover of \(X\), collection of open sets

\[ U = \{ U_\alpha \} \quad \alpha \in A \rightarrow \text{ finite index set} \]

s.t.

\[ \bigcup_{\alpha \in A} U_\alpha = X \quad \text{(assume } U_\alpha \text{ is path-connected)} \]

\( \bigcup_{\alpha \in A} (U_\alpha \cap X) = X \)

**Example:**

\[ X = S^1 \subset \mathbb{R}^2 \quad \text{(cycle)} \]

\[ A = \{ 1, 2, \ldots, 10 \} \]

\[ U \bigcup_{\alpha \in A} (U_\alpha \cap X) = X \]

\[ \rightarrow \text{ Alternatively, } \quad X \subset \bigcup_{\alpha \in A} U_\alpha \]

**Def.:** Nerve of a cover: Given a finite cover \( U = \{ U_\alpha \} \quad \alpha \in A \) of \( X \), the nerve of cover \( U \) is the simplicial complex \( N(U) \) whose vertex set is the index set \( A \) and a subset \( \{ \alpha_0, \alpha_1, \ldots, \alpha_k \} \subseteq A \) spans a \( k \)-simplex in \( N(U) \) iff

\[ U_{\alpha_0} \cap U_{\alpha_1} \cap \cdots \cap U_{\alpha_k} \neq \emptyset \]

(intersection of open sets indexed by \( \alpha_0, \ldots, \alpha_k \) is not empty)

**Čech complex:** Nerve complex of a collection of open balls
Def: [Mapper] Let $X, Z$ be topological spaces and $f: X \to Z$ be a well-behaved continuous function ($Z = \mathbb{R}$ or $\mathbb{R}^2$) for example.

Let $U = \bigcup_{x \in X} U_x^3$ be a finite open cover of $Z$.

The mapper (mapper construction) is defined to be the nerve complex of the pull-back cover $M(u, f) = N(f^*(u))$.

Def: pull-back cover. Let $f: X \to Z$ as above, and $Z$ be equipped with cover $U = \bigcup_{x \in A} U_x^3$.

The sets $\bigcup f^{-1}(U_x^3)$ form an open cover of $X$.

Consider the path-connected components of $f^{-1}(U_x^3)$:

$$f^{-1}(U_x^3) = \bigcup_{i=1}^n V_{x,i};$$

Let $D$ be path-connected components of $f^*(U)$.

$f^*(U)$ is the cover of $X$ obtained from cover $U$ of $Z$.

It is the pull-back cover of $X$ (induced by $U$ via $f^*$).

---

Suppose $U_1 = (0, 3)$, $U_2 = (2, 7)$, $U_3 = (6, 10)$, $U = \bigcup_{i=1}^3 U_i$.

For $f: X \to Z$, $Z \in \mathbb{R}^2$.
Def: Well-behaved function $f$: for every path-connected open set $U \subseteq \mathbb{R}^1$, $f^{-1}(U)$ has finitely many path-connected components.

E.g., piecewise linear real-valued function on a finite S.C.

Mappers gives a "soft clustering" of domain (clusters overlap) but it also gives a relationship between clusters and it can preserve topological structure (depending on the covers we define).