Connectivity-Optimized Representation Learning with Persistent Homology

Amin Dhaou, Yassine Kamri, Patrick Saux

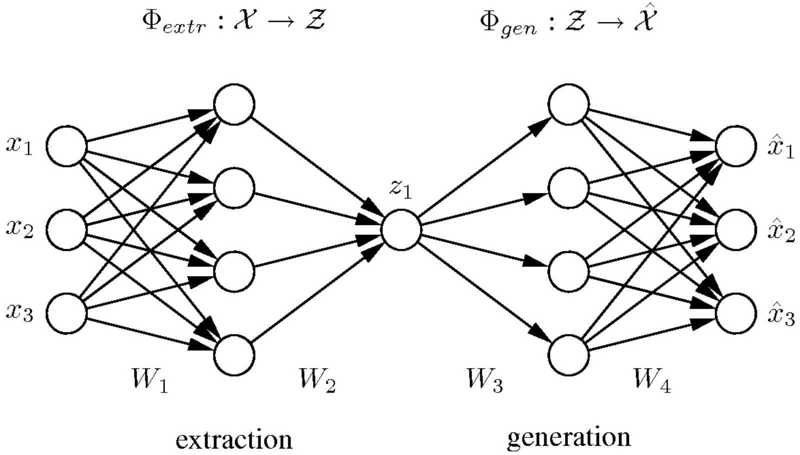
Topological Data Analysis

January 16th 2020

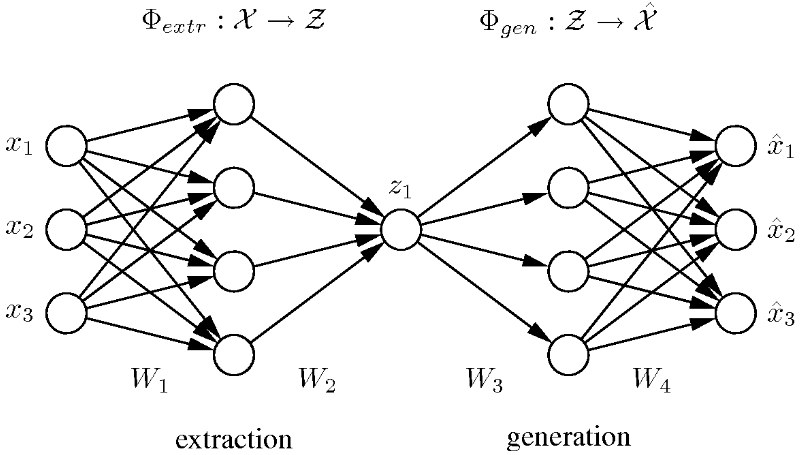
Outline

Code available at :
`https://github.com/sauxpa/connectivity_loss`

Autoencoder



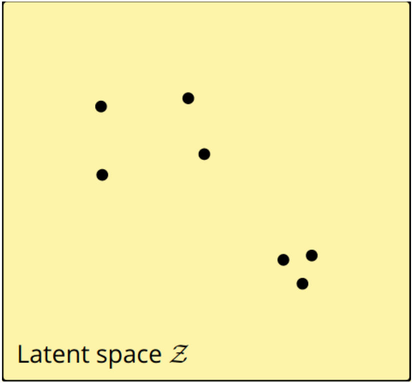
Autoencoder



Possible regularization of $\mathcal{X}$ and $\hat{\mathcal{X}}$, but no direct control of $\mathcal{Z}$.

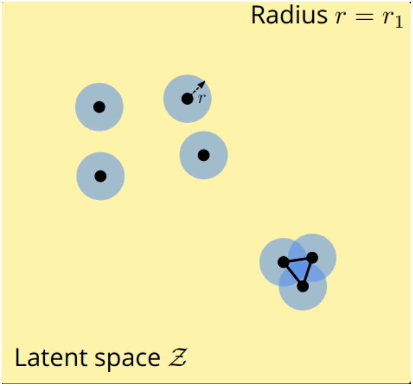
Vietoris-Rips filtration

https://icml.cc/media/Slides/icml/2019/grandball(12-16-00)-12-16-25-4491-connectivity-op.pdf



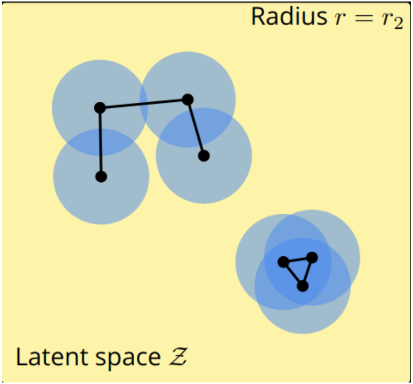
Vietoris-Rips filtration

https://icml.cc/media/Slides/icml/2019/grandball(12-16-00)-12-16-25-4491-connectivity-op.pdf



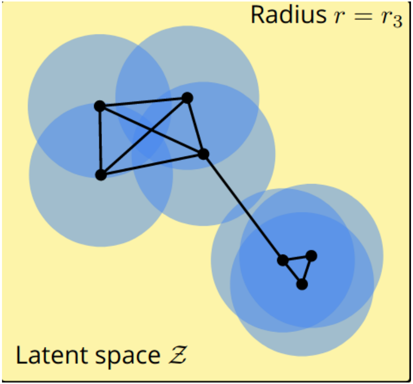
Vietoris-Rips filtration

https://icml.cc/media/Slides/icml/2019/grandball(12-16-00)-12-16-25-4491-connectivity-op.pdf



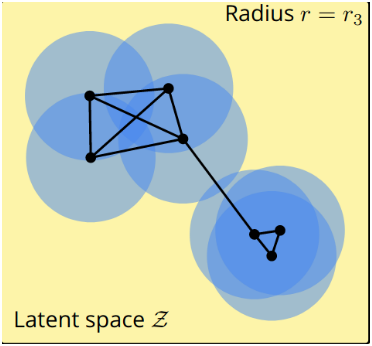
Vietoris-Rips filtration

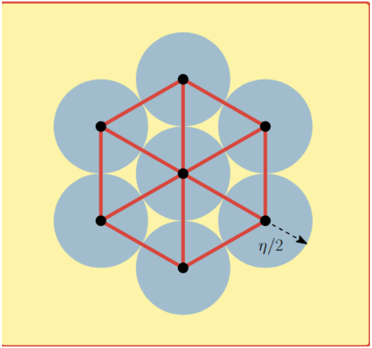
https://icml.cc/media/Slides/icml/2019/grandball(12-16-00)-12-16-25-4491-connectivity-op.pdf



Vietoris-Rips filtration

https://icml.cc/media/Slides/icml/2019/grandball(12-16-00)-12-16-25-4491-connectivity-op.pdf





Connectivity loss

Definition

Definition (Connectivity loss)

Let $(\mathbb{R}^n, d)$ a metric space, $S = \{z_1, \ldots, z_b\} \subset \mathbb{R}^n$, $(\varepsilon_k)_{k=1}^{M}$ the sequence of increasing pairwise distances between points in S and $\eta > 0$:

1. $\emptyset = \mathcal{V}_0 \subset \mathcal{V}_{\frac{\varepsilon_1}{2}} \subset \cdots \subset \mathcal{V}_{\frac{\varepsilon_M}{2}}$ Vietoris-Rips filtration,
2. $\dagger(S) = \{t > 0 | (0, \frac{\varepsilon_t}{2}) \in barcode(S)\}$ 0-homology death times,
3. $\mathcal{L}_\eta(S) = \sum_{t \in \dagger(S)} |\eta - \varepsilon_t|$ connectivity loss.

Connectivity loss

Definition

Definition (Connectivity loss)

Let $(\mathbb{R}^n, d)$ a metric space, $S = \{z_1, \ldots, z_b\} \subset \mathbb{R}^n$, $(\varepsilon_k)_{k=1}^M$ the sequence of increasing pairwise distances between points in S and $\eta > 0$:

1. $\emptyset = \mathcal{V}_0 \subset \mathcal{V}_{\frac{\varepsilon_1}{2}} \subset \cdots \subset \mathcal{V}_{\frac{\varepsilon_M}{2}}$ Vietoris-Rips filtration,
2. $\dagger(S) = \{t > 0 | (0, \frac{\varepsilon_t}{2}) \in barcode(S)\}$ 0-homology death times,
3. $\mathcal{L}_\eta(S) = \sum_{t \in \dagger(S)} |\eta - \varepsilon_t|$ connectivity loss.

Intuition: the connectivity loss penalizes configurations that diverge from regular arrangements.

Connectivity loss

Differentiability

Theorem

Define the indicator :

$$\mathbb{1}_{ij}(z_1, \dots, z_b) = \begin{cases} 1, & \textit{if } \exists t \in \dagger(S) : \varepsilon_t = d(z_i, z_j), \\ 0, & \textit{otherwise.} \end{cases}$$

Then:

- (i) $\mathcal{L}_\eta(S) = \sum_{i,j} |\eta - d(z_i, z_j)| \mathbb{1}_{ij}(z_1, \dots, z_b)$,
- (ii) $\mathcal{L}_\eta(S)$ *is (sub)-differentiable w.r.t coordinates* z_{uv} *and* $\partial_{z_{uv}} \mathcal{L}_\eta(S) = \sum_{i,j} \partial_{z_{uv}} |\eta - d(z_i, z_j)| \mathbb{1}_{ij}(z_1, \dots, z_b)$.

Connectivity loss

Differentiability

Theorem

Define the indicator :

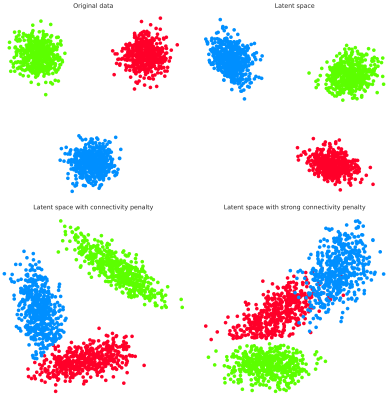
$$\mathbb{1}_{ij}(z_1, \ldots, z_b) = \begin{cases} 1, & \textit{if } \exists t \in \dagger(S) : \varepsilon_t = d(z_i, z_j), \\ 0, & \textit{otherwise.} \end{cases}$$

Then:

1. $\mathcal{L}_\eta(S) = \sum_{i,j} |\eta - d(z_i, z_j)| \mathbb{1}_{ij}(z_1, \ldots, z_b)$,
2. $\mathcal{L}_\eta(S)$ *is (sub)-differentiable w.r.t coordinates* z_{uv} *and* $\partial_{z_{uv}} \mathcal{L}_\eta(S) = \sum_{i,j} \partial_{z_{uv}} |\eta - d(z_i, z_j)| \mathbb{1}_{ij}(z_1, \ldots, z_b)$.

... can backpropagate!

Geometric Density



0-homology in high dimension

Contrast degeneracy (Aggarwal et al.)

Theorem (Aggarwal)

Let $Z_1^d, \ldots, Z_b^d \in \mathbb{R}^d$ *sampled uniformly in the hypercube* $[0,1]^d$*, define the d-contrast as* $C_d(Z_1^d, \ldots, Z_b^d) = \frac{\max_i \|Z_i^d\| - \min_i \|Z_i^d\|}{\min_i \|Z_i^d\|}$*. Then:*

$$C_d(Z_1^d, \ldots, Z_b^d) \xrightarrow[d \to \infty]{\mathbb{P}} 0.$$

0-homology in high dimension

Contrast degeneracy (Aggarwal et al.)

Theorem (Aggarwal)

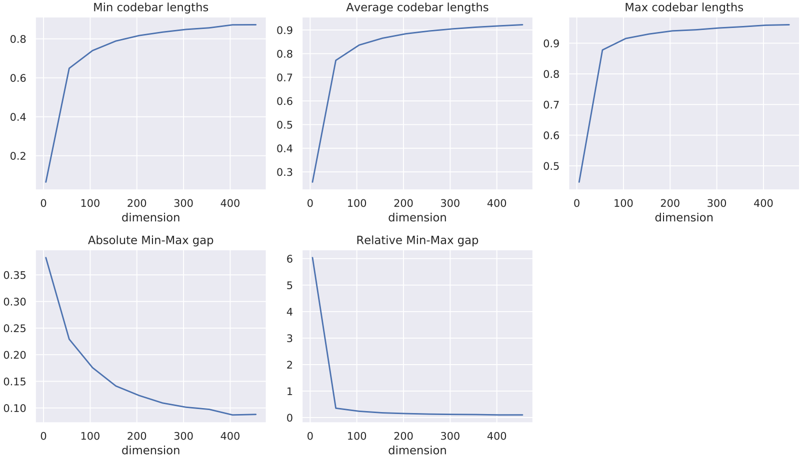
Let $Z_1^d, \ldots, Z_b^d \in \mathbb{R}^d$ *sampled uniformly in the hypercube* $[0,1]^d$, *define the d-contrast as* $C_d(Z_1^d, \ldots, Z_b^d) = \frac{\max_i \|Z_i^d\| - \min_i \|Z_i^d\|}{\min_i \|Z_i^d\|}$. *Then:*

$$C_d(Z_1^d, \ldots, Z_b^d) \xrightarrow[d \to \infty]{\mathbb{P}} 0.$$

The higher the dimension, the less discriminant the distance.

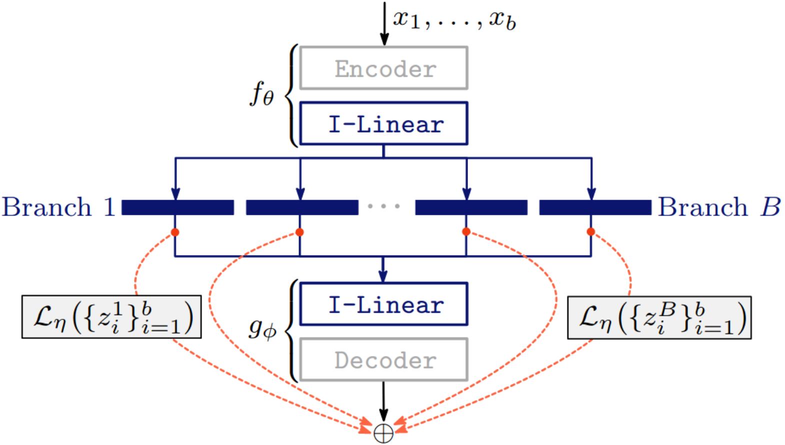
0-homology in high dimension

Empirical evidences



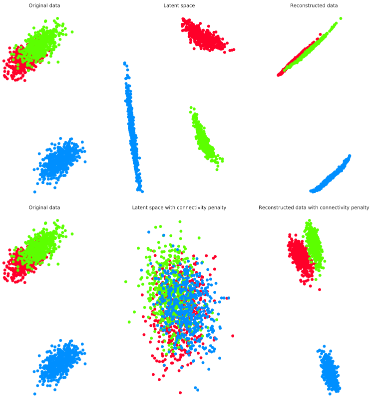
0-homology in high dimension

Branching architecture: https://icml.cc/media/Slides/icml/2019/grandball(12-16-00)-12-16-25-4491-connectivity-op.pdf



Experiments

Gaussian mixtures in high dimension



Experiments

CIFAR-10 (in-sample)

Experiments

CIFAR-10 (out-of-sample)

Conclusion

1. Theoretical guarantees of densification
2. Regularization properties (better generalization)
3. Possible extensions:
    1. Trainable η
    2. Higher homology
    3. Other latent space target, connexion with persistence diagram vector features...

Conclusion

THANKS!

0-homology in high dimension

Penrose theorem

Theorem (Penrose)

Let $\epsilon > 0$, G the graph constructed from b points $Z_1, \ldots, Z_b$ sampled uniformly in the hypercube $[0,1]^d$ with an edge between Z_i and Z_j if and only if $d(Z_i, Z_j) \leq \varepsilon$. Connectedness of G exhibits a phase transition with threshold:

$$\varepsilon \sim \left(\frac{b}{\log b}\right)^{\frac{1}{d}}.$$

0-homology in high dimension

Penrose theorem

Theorem (Penrose)

Let $\epsilon > 0$, G the graph constructed from b points $Z_1, \dots, Z_b$ sampled uniformly in the hypercube $[0,1]^d$ with an edge between Z_i and Z_j if and only if $d(Z_i, Z_j) \leq \varepsilon$. Connectedness of G exhibits a phase transition with threshold:

$$\varepsilon \sim \left(\frac{b}{\log b}\right)^{\frac{1}{d}}.$$

The higher the dimension, the "harder" it is to connect random geometric graphs.

Experiments
Gaussian mixtures with trainable η