TopoSampler: A topology constrained noise sampling in GANs

Adrish Dey, Sayantan Das

Netaji Subhash Engineering College, Institute of Engineering and Management

Objectives

- Introduce a topological perspective towards disconnected manifold learning in GANs
- Demonstrate the importance of prior-space topology in disconnected manifold learning.
- Run initial experiments on a novel noise module for learning topologically constrained priors.
- Run initial experiments on a homeomorphism constrained generator.

Introduction

Learning distribution with disconnected manifold is a challenging problem. The distribution learned by a regular vanilla GAN, with unimanifolded-prior, creates a global cover over the disconnected components. This produces samples, which are unrealistic interpolation between two completely distinct classes.

Motivation for exploring topological perspective towards disconnected manifold learning:

- Disconnectedness in manifold is topologically represented as 0^{th} dimensional holes in space. This makes it interesting to study a generalized form of manifold disconnectedness in terms of d-dimensional holes.
- current method are either resource intensive (multi-generator) or have lower overall precision.(rejection sampling)

Proposed Solution

We found empirical evidence on the important of prior space topology in learning disconnected posteriors. Based on that we divide the learning problem in two parts:

- **Prior Topology Optimization**: We explicitly introduce d-dimensional holes in the unimanifolded prior space, by regularizing it to be homologically similar to the data space
- **Homeomorphic density estimation**: Next, we learn an isometric map from the learned prior to the posterior space. This ensures, the d-dimensional holes in the prior space, are preserved during GAN training.

Method Overview

- By leveraging a Neural Network, we learn a mapping from a gaussian distribution to a latent space. Using results from [1] we create a synthetic prior space, with a sample space regularized to be topologically similar to the data space.
- We use a lispchitz-constrained GAN to learn a homeomorphism between the synthetic prior space and the target data space (i.e., the posterior space). The weights of the generator are singular-value normalized to ensure 1-Lipschitz continuity. Along with 1-Lipschitz learnable layers, usage of 1-Lipschitz activations in the generator ensures $L \leq 1$ throughout, thus explicitly establishing an isometry, thereby homeomorphism. [2]

Optimization Objective

Let $\mathcal{X} = \{x_0, x_1, \dots\}$ with $x_i \in \mathbb{R}^d$ denote the original data samples, which we consider to be samples from an underlying manifold $\mathbb{M}_{\mathcal{X}}$, and probability distribution $P(\mathcal{X})$. Let the generator of the proposed GAN be represented as $\mathcal{G}_\theta$ and the discriminator as $\mathcal{D}_\phi$. Similarly, a parametric function N_ψ with a latent sample space $\mathfrak{N}$, is used for mapping $\mathcal{N}(0,1) \to P(\mathfrak{N})$, whose sample space manifold, $\mathbb{M}_\eta$, is explicitly regularized to be topologically similar to $\mathbb{M}_{\mathcal{X}}$ This topological similarity is explicitly induced using signature loss $\mathcal{L}(\mathcal{X}, \mathfrak{N})$ defined as:

$$\mathcal{L}(\mathfrak{N}, \mathcal{X}) = \frac{1}{2} \left\| D_{\mathcal{X}}[\Pi_{\mathcal{X}}] - D_{\mathfrak{N}}[\Pi_{\mathcal{X}}] \right\|^2 + \frac{1}{2} \left\| D_{\mathcal{X}}[\Pi_{\mathfrak{N}}] - D_{\mathfrak{N}}[\Pi_{\mathfrak{N}}] \right\|^2$$

Where $D_{\mathcal{X}}$ and $D_{\mathfrak{N}}$ denotes the pairwise distances between sample sets $\mathcal{X}$ and $\mathfrak{N}$ respectively. $\Pi_{\mathcal{X}}$ and $\Pi_{\mathfrak{N}}$ are the indices of topologically significant simplices found by Vietoris–Rips Filtration $\mathfrak{R}_\epsilon(\mathcal{X})$ and $\mathfrak{R}_\epsilon(\mathfrak{N})$ of sample spaces $\mathcal{X}$ and $\mathfrak{N}$.

In adversarial optimization, the discriminator objective remains the same, Whereas the generator loss is appended with $\mathcal{L}$ as a regularization term.

$$\mathbb{E}_{z' \sim P(\mathfrak{N})} \left[\ln(1 - \mathcal{D}_\phi(\mathcal{G}_\theta(z'))) \right] + \left[\frac{1}{B} \sum \mathcal{L}(N_\psi(z), x_r) \right]_{z \sim \mathcal{N}(0,1), x_r \sim P(\mathcal{X})}$$

Experiments

The following observations are made in our experiments.

- increasing the dimension of prior space, reduces the quality of samples generated by a GAN.
- Neural Networks are inherently not suitable are introducing holes in the manifold. Demonstrated by the "stretching effect" of a manifold when topological regularization is performed. (See figure 1(a))
- Lipschitz constraining the generator although makes the training highly unstable, does have a homeomorphic effect with respect to the prior. (See figure 1(b))

(a) (b)

Figure 1(a), the sparsely spread samples between two distribution demonstrates the "stretching effect" of topo-regularization. Figure 1(b), GAN learns a hole preserving map between prior and posterior. Here disconnected prior was used to learn a disconnected posterior.

For more information please refer to the paper.

Conclusion

- This work studies the problem of learning disconnected sample space manifolds in GANs by topologically aligning the prior space to the original data space.
- We introduce a persistent homology perspective towards augmenting the prior distribution to stay in the same homology class as that of our unknown data manifold.

References

[1] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt.
Topological autoencoders, 2020.

[2] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida.
Spectral normalization for generative adversarial networks.
In *International Conference on Learning Representations*, 2018.

[3] Mahyar Khayatkhoei, Ahmed Elgammal, and Maneesh Singh.
Disconnected manifold learning for generative adversarial networks, 2019.

Acknowledgements

We would like to extend our deepest gratitude to Dr. Bastian Rieck, D-BSSE, ETH Zurich, for his unending support during experimentation and documentation of the project

Contact Information

Adrish Dey

- Web: https://member.acm.org/~adrishd
- Email: adrishd.cse2017@nsec.ac.in

Sayantan Das

- Web: https://github.com/ucalyptus
- Email: sdas.codes@gmail.com