Benchmarking Graph Convolutional Networks Performance with Mixed-Product Space Embedding

Fangchen Li **Lingjue Xie** **Zelin Li**

{fcli1997, ljxie, jameszli}@stanford.edu

Motivation

Non-Euclidean spaces have gained popularity due to their ability to better capture the topology of data. Embeddings in a product space further improve data representations quality by providing heterogeneous curvature suitable for a wide variety of structures.

Objective

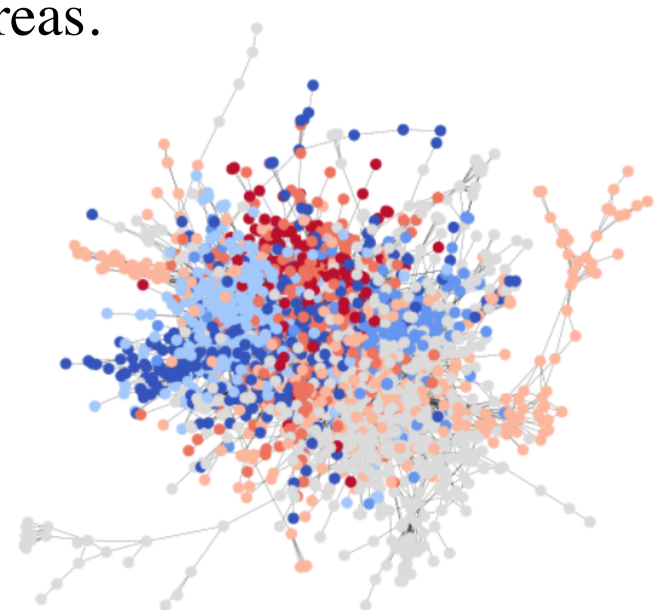
- Extending Graph Convolutional Networks (GCNs) to mixed-product space
- Benchmarking performances to provide heuristics on choosing the best mixture space

Dataset

We used four open datasets for this project.

CORA and PUBMED describe citation network where nodes represent scientific papers, edges are citations between them, and node labels are academic areas.

AIRPORT is a flight network dataset where nodes represent airports, edges represent the airline routes, and node labels are population of the country where the airport belongs.



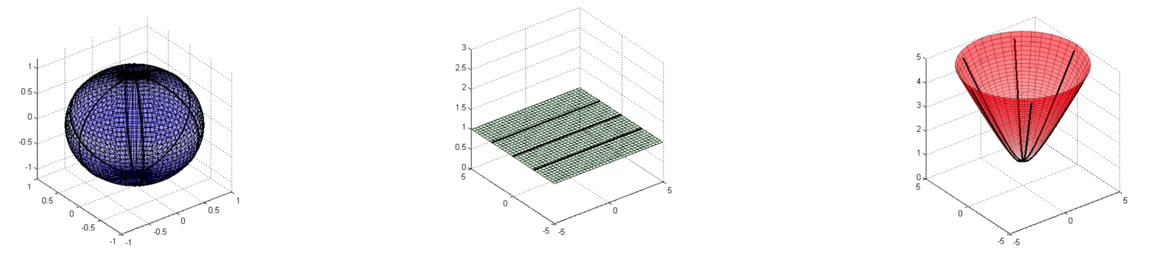
CORA dataset

DISEASE is a simulated SIR disease spreading model, where the label of a node is whether the node was infected or not.

Methodology

Mixed-Product Space

We are primarily interested in using the following three spaces in creating mixed-product spaces.



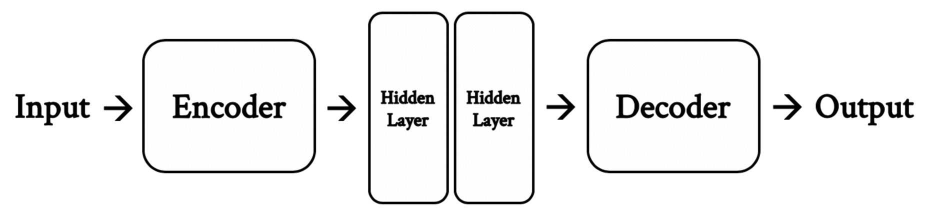
Sphere $\mathbb{S}^2$ **Euclidean $\mathbb{E}^2$** **Hyperboloid $\mathbb{H}^2$**

We tested the following feature dimensions for mixed-product space: [4, 8, 16, 32, 64], with each dimension being distributed according to some ratios combining the three spaces.

<u>Formation of Mixed-Product Space</u>

$$\mathcal{P} = \mathbb{S}^{s_1} \times \mathbb{S}^{s_2} \times \cdots \times \mathbb{S}^{s_m} \times \mathbb{H}^{h_1} \times \mathbb{H}^{h_2} \times \cdots \times \mathbb{H}^{h_n} \times \mathbb{E}^{e}$$

Graph Convolutional Networks (GCNs)



GCN Structure

1. **Input:** Node features and adjacency matrix of graphs
2. **Encoder:** Euclidean node features projected to the product space
3. **Hidden Layers:**

$$h_i^{l,\mathcal{P}} = (W^l \otimes x_i^{l-1,\mathcal{H}}) \oplus b^l \quad \textbf{Transform}$$

$$y_i^{l,\mathcal{P}} = AGG(h_i^{l,\mathcal{P}}) = \exp_{h_i^{\mathcal{P}}}\Big(\sum_{j\in\mathcal{N}(i)} \log_{h_i^{l,\mathcal{P}}}(h_j^{l,\mathcal{P}})\Big) \quad \textbf{Aggregation}$$

$$x_i^{l,\mathcal{P}} = \sigma^{\otimes}(y_i^{l,\mathcal{P}}) = \exp_o(\sigma(Log_o(y_i^{l,\mathcal{P}}))) \quad \textbf{Activation}$$

4. **Decoder & Output:**

<u>Linkage Prediction (LP)</u>

- Use Fermi-Dirac decoder to compute probabilities for linkages
- Minimize loss using cross-entropy with negative sampling

$$p((i,j) \in \mathcal{E} | x_i^{L,\mathcal{P}}, x_j^{L,\mathcal{P}}) = \frac{1}{1 + \exp\left(\frac{d_{\mathcal{P}}(x_i^{L,\mathcal{P}}, x_j^{L,\mathcal{P}})^2 - r}{t}\right)}$$

Predictive Probability of New Edge (LP)

<u>Node Classification (NC)</u>

- Project the output from product space into Euclidean space
- Conduct multinomial logistic regression

Performance & Discussion

	AIRPORT		CORA		DISEASE	
Dimension	HGCN($\mathbb{H}^{dim}$)	Product Space	HGCN($\mathbb{H}^{dim}$)	Product Space	HGCN($\mathbb{H}^{dim}$)	Product Space
4	0.8754	0.8822 ($\mathbb{E}^4$)	0.8638	0.8553 ($\mathbb{S}^2 \times \mathbb{S}^2$)	0.5010	0.5214 ($\mathbb{S}^2 \times \mathbb{S}^2$)
8	0.8976	0.9086 ($\mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{H}^4$)	0.9107	0.8672 ($\mathbb{S}^4 \times \mathbb{S}^4$)	0.3965	0.5363 ($\mathbb{S}^4 \times \mathbb{S}^4$)
16	0.8891	**0.9286** ($\mathbb{S}^8 \times \mathbb{H}^8$)	0.9187	0.8914 ($\mathbb{H}^8 \times \mathbb{H}^8$)	0.5746	0.5684 ($\mathbb{H}^8 \times \mathbb{H}^8$)
32	0.9070	0.9251 ($\mathbb{S}^{16} \times \mathbb{H}^{16}$)	0.9351	0.9065 ($\mathbb{S}^8 \times \mathbb{S}^8 \times \mathbb{H}^{16}$)	0.6036	0.5804 ($\mathbb{H}^{16} \times \mathbb{H}^{16}$)
64	0.9203	0.9275 ($\mathbb{S}^{32} \times \mathbb{H}^{32}$)	**0.9371**	0.9176 ($\mathbb{S}^{32} \times \mathbb{H}^{32}$)	**0.6375**	0.6000 ($\mathbb{E}^{64}$)

Table 1: Linkage Prediction(LP) ROC - AUC

	AIRPORT		CORA		PUBMED	
Dimension	HGCN($\mathbb{H}^{dim}$)	Product Space	HGCN($\mathbb{H}^{dim}$)	Product Space	HGCN($\mathbb{H}^{dim}$)	Product Space
4	0.4771	0.4905 ($\mathbb{S}^2 \times \mathbb{S}^2$)	0.7050	0.7050 ($\mathbb{E}^4$)	0.7600	0.7640 ($\mathbb{S}^2 \times \mathbb{S}^2$)
8	0.5210	0.5553 ($\mathbb{S}^4 \times \mathbb{S}^4$)	0.7540	0.7540 ($\mathbb{S}^8$)	0.7690	0.7690 ($\mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{H}^2$)
16	0.6088	0.6088 ($\mathbb{S}^{16}$)	0.7890	0.7890 ($\mathbb{S}^{16}$)	0.7510	**0.7740** ($\mathbb{S}^4 \times \mathbb{S}^4 \times \mathbb{H}^4 \times \mathbb{H}^4$)
32	**0.6183**	**0.6183** ($\mathbb{S}^{32}$)	0.7820	0.7820 ($\mathbb{S}^{32}$)	0.7690	0.7720 ($\mathbb{S}^{16} \times \mathbb{S}^{16}$)
64	0.6145	**0.6183** ($\mathbb{S}^{32} \times \mathbb{H}^{16} \times \mathbb{H}^{16}$)	0.7780	**0.7920** ($\mathbb{S}^{32} \times \mathbb{S}^{32}$)	0.7640	0.7700 ($\mathbb{S}^{32} \times \mathbb{H}^{16} \times \mathbb{H}^{16}$)

Table 2: Node Classification(NC) F1-Score

Linkage Prediction: Performance of mixed-product spaces strongly depends on the topology of the dataset. Mixed-product space performs better on datasets with spherical structures whereas the baseline HGCN has better performance on hierarchical data.

Node Classification: For all datasets, product spaces either outperform or have same output as the HGCN model, suggesting a better alternative.

Future Work

1. Explore more possibilities in mixed-product space, such as including non-constant curvature spaces as our component
2. Learn curvature as the GCN model's parameter for each hidden layer and add attention mechanism to our model
3. Test additional datasets to further understand the relationships between mixed-product spaces and structures of datasets

References

[1] Chami, Ines, et al. "Hyperbolic Graph Convolutional Neural Networks." ArXiv:1910.12933 [Cs, Stat], Oct. 2019. arXiv.org, http://arxiv.org/abs/1910.12933.

[2] Gu, A., Sala, F., Gunel, B., & Ré, C. (2018). Learning Mixed-Curvature Representations in Product Spaces.

[3] G. Bachmann, G. Bécigneul, and O.-E. Ganea, "Constant Curvature Graph Convolutional Networks," 12-Nov-2019. [Online]. Available: https://arxiv.org/abs/1911.05076.

Stanford Statistics