Low-dimensional Knowledge Graph Embedding based on Extended Poincaré Ball: Preliminary Results ⋆

Xingchen Zhou, Zhe Pan, and Peng Wang(✉)

School of Computer Science and Engineering, Southeast University, Nanjing, China
{xczhou_2021, zpan, pwang}@seu.edu.cn

Abstract. Most existing knowledge graph embedding (KGE) methods are built on Euclidean space, which are difficult to handle hierarchical structures. While hyperbolic embedding methods have shown the promise of high fidelity and concise representation for hierarchical data, logical patterns in knowledge graphs (KGs) are not considered well in current methods. To address this problem, we propose a novel KGE model with extended Poincaré Ball and polar coordinate system to capture hierarchical structures. It first uses the tangent space and exponential transformation to initialize and map the corresponding vectors to the Poincaré Ball in hyperbolic space. Then, to solve the boundary conditions, the Poincaré Ball boundary is stretched and zoomed by expanding the modulus length. Moreover, it is optimized by using polar coordinate and changing operators in the extended Poincaré Ball. Experimental results of link prediction on WN18RR and FB15k-237 datasets show that our model outperforms state-of-the-art baselines.

Keywords: Knowledge graph embedding · Extended Poincaré Ball · Hyperbolic space

1 Introduction

Since knowledge graphs (KGs) are usually incomplete, predicting missing links in KGs via knowledge graph embedding (KGE) into vector spaces becomes more and more important. Hierarchical structures are common in KGs and used to manage the relations and concepts. However, existing KGE methods often encounter challenges dealing with hierarchical structures, because it is notably difficult for models built on Euclidean space to preserve hierarchical structures.

Recent works proposed hyperbolic representation learning [2, 6]. In KGs, hierarchical relationships between entities can be approximated as a tree structure,

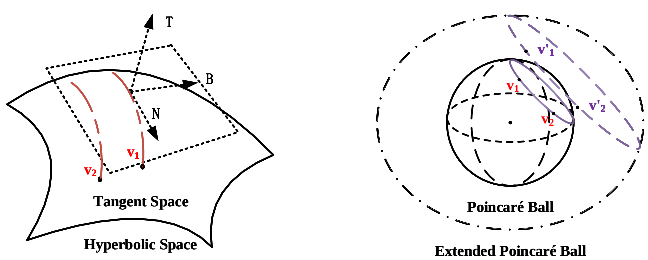
Corresponding author: Peng Wang (pwang@seu.edu.cn). The work is supported by All-Army Common Information System Equipment Pre-Research Project (No. 31514020501, No. 31514020503)

while the number of entities in each layer increases exponentially with depth of tree increasing. Such a knowledge structure can be well represented with the Poincaré Ball [9]. However, even most hyperbolic KGE models employ Poincaré Ball to embed the structures, they still suffer from limitations of restricted capacity and floating-point precision when majority of entities are embedded near by the boundary of Poincaré Ball due to long-tail distribution.

To address these issues, we propose a novel hyperbolic knowledge embedding method, which employs the extended Poincaré Ball for KGE and captures hierarchical structures with polar coordinate system [5].

2 Model

The principle of our model is shown in Fig. 1. In order to learn hierarchical hyperbolic embeddings to represent logical patterns such as symmetry and anti-symmetry while preserving latent hierarchies, our model uses polar coordinate to encode logical patterns with hierarchies in hyperbolic space as shown in Fig. 1(a). Due to the advantages of Poincaré Ball for gradient optimization [1], our model first initials embedding in Poincare Ball. However, the capacity of Poincaré Ball model is restricted by floating-point precision when majority of points locate near the boundary due to long-tail distribution. To release this limitation, our method expands the boundary into infinite to increase capacity of model and adjust some operators to align with Euclidean geometry, which redefines the distance $d_{\mathcal{B}}(u, v)$ and can be shown in Fig. 1(b).



(a) Hyperbolic Space. (b) Extended Poincaré Ball.

Fig. 1. KGE based on Hyperbolic space and KGE based on extended Poincaré ball.

In Poincaré Ball, the whole space is symmetric along the center but the apparent Euclidean distance from the origin to any point is not equal to the hyperbolic distance [5]. To make the apparent distance consistent with the actual hyperbolic distance, we establish a new model called the extended Poincaré Ball to ensure the distance from any point to the center is equal to which in hyperbolic space. Suppose that the polar coordinates of any point in the original coordinate system (Poincaré disk) is(r, θ), and that in the new space is $(2tanh^{-1}r, \theta)$. In this way, the radius of ball space is infinite. Therefore, points near the boundary are extremely compressed in Poincaré Ball while there is no such problem in the extended one. Meanwhile, it can be proved that Hyperbolic Cosine Theorem still holds for the operators in extended Poincaré Ball: $cosh(c) = cosh(a)cosh(b) - sinh(a)sinh(b)\cos\gamma$ $(a, b, c$ stand for the geodesic distance of triangle and γ stands for the angle between $a, b)$. Extended Poincaré Ball and Poincaré Ball share the same distance form when calculated by cosine theorem as well.

Furthermore, inspired by the Hyperbolic Cosine Theorem, in which the hyperbolic distance can be composed of modulus and angle, we use polar coordinates to embed KGs into the extended Poincaré Ball. The score function can be formed as two parts: polar radius $(\mathbf{h}, \mathbf{r}, \mathbf{t})$ and polar angle $(\theta_h, \theta_r, \theta_t)$ as follow:

$$d_{\mathcal{B}} = \alpha d_{r\mathcal{B}} + \beta d_{\theta\mathcal{B}} \tag{1}$$

where α, β is the weights to be learned. The whole function shares the similar way as works proposed by Federico [7] which does not satisfy Cauchy inequality.

It is worth noting that the radius part plays an essential role in levels of entities in extended Poincaré Ball, and the angle aims to distinguish entities in the same level. Therefore, we formulate the polar radius with Möbius addition and multiplication as follow:

$$d_{r\mathcal{B}} = \|2\tanh^{-1}((\mathbf{R} \otimes_c \mathbf{h}) \oplus_c -(\mathbf{r} \oplus_c \mathbf{t}))\|^2 \tag{2}$$

where $\mathbf{h}, \mathbf{r}, \mathbf{t}$ stand for hyperbolic embeddings of head entity, relation, and tail entity, respectively. $\mathbf{R}$ stands for relation matrix in hyperbolic space inspired by MuRP [2]. As stated in the property of extended Poincaré Ball, we classify the embedding levels of different entities by Euclidean Norm.

Considering convergence and efficiency, we can simplify and obtain the polar angle as: $\Delta\theta = \pi - |\pi - |\theta - \theta'||$. Consequently, a point $x_{\mathcal{B}}$ in polar coordinate system can be calculated in TransE form [4] as:

$$d_{\theta\mathcal{B}} = \|(\theta_h + \theta_r - \theta_t) mod 2\pi\| \tag{3}$$

From another perspective, angle parts can be replaced with radius parts by Cosine Theorem in hyperbolic space. However, to better capture complex relation such as symmetry, anti-symmetry, inversion and composition, it is necessary to utilize extra angle part for downstream tasks like link predictions. On the other hand, the introduction of angle part can simulate the rotation in RotatE [8]. Theoretically, any algebraic system hold the fundamental properties of congruence can be used as angle part in the model when embedding complex relations. Take angles as an example, suppose that a relation $\theta_r \in [0, 2\pi)$ is close to π, then a symmetric relation can be formed as $(\theta_h + \theta_r + \theta_r)\, mod\, 2\pi = \theta_h\, mod\, 2\pi$ with arbitrary θ_h and $\neq$ for asymmetric relations.

Since the Poincaré Ball has a Riemannian manifold structure, we optimize radius parameters with stochastic Riemannian optimization methods such as RSGD or RSVRG [3]. Let ∇E denote the Euclidean gradient of $L(P)$. Using

RSGD, the Riemannian gradient can be computed as $\nabla_R = \frac{\left(1-\|P_t\|^2\right)^2}{4}\nabla_E$. In summary, the full update for a single embedding is calculated by:

$$\boldsymbol{P}_{t+1} = \boldsymbol{P}_t - \eta_t \frac{\left(1 - \|\boldsymbol{P}_t\|^2\right)^2}{4}\nabla_E \tag{4}$$

where η denotes the learning rate.

According to the isometric projection of Poincaré Ball, the angle part can be optimized by Euclidean optimization methods such as SGD or Adam.

To train the model, we use the negative sampling loss functions with self-adversarial training [8].

$$\begin{aligned} s = & -\log\left(\sigma\left(\lambda - d_{\mathcal{B}}(\mathbf{h}, \mathbf{r}, \mathbf{t})\right)\right) \\ & -\sum_{i=1}^{n} p\left(\mathbf{h}_\mathbf{i}', \mathbf{r}, \mathbf{t}_\mathbf{i}'\right) \log\left(\sigma\left(d_{\mathcal{B}}\left(\mathbf{h}_\mathbf{i}', \mathbf{r}, \mathbf{t}_\mathbf{i}'\right) - \lambda\right)\right) \end{aligned} \tag{5}$$

where λ is margin.

For negative samples,

$$p\left(\mathbf{h}_\mathbf{j}', \mathbf{r}, \mathbf{t}_\mathbf{j}' \mid \{(\mathbf{h}_\mathbf{k}, \mathbf{r}_\mathbf{k}, \mathbf{t}_\mathbf{k})\}\right) = \frac{e^{yf\left(\mathbf{h}_\mathbf{j}', \mathbf{t}_\mathbf{j}\right)}}{\sum_{k=1}^{size} e^{yf\left(\mathbf{h}_\mathbf{k}', \mathbf{t}_\mathbf{k}'\right)}} \tag{6}$$

where p is the probability distribution of sampling negative triples, and α is the temperature of sampling.

3 Experiments

Table 1. Link prediction results on WN18RR and FB15k-237.

Models	WN18RR			FB15k-237		
	MRR	H@1	H@10	MRR	H@1	H@10
TransE [4]	.226	.043	.501	.294	.204	.465
DisMult [10]	.437	.397	.490	.241	.155	.419
RotatE [8]	.476	.428	.571	.338	.241	.533
HyperKG [6]	.41	-	.50	.28	-	.45
MuRP [2]	.481	.440	.566	.323	.235	.501
Our Model	**.488**	**.448**	**.570**	**.340**	**.243**	**.541**

To evaluate our approach, we choose the widely used KG datasets: WN18RR and FB15K-237. The evaluation metrics are: (1) mean reciprocal rank (MRR), which measures the mean of inverse ranks assigned to correct entities; and (2) hits at K (H@K, $K \in 1, 10$), which measures the proportion of correct triples among the top-K predicted triples. Table 1 summarizes the experimental results of different models on the task of KG link prediction. It can be seen that our method achieves the state-of-the-art results on both datasets, which demonstrates the efficiency of our model.

References

1. Abramowicz, M.A., Bengtsson, I., Karas, V., Rosquist, K.: Poincaré ball embeddings of the optical geometry. Classical and Quantum Gravity **19**(15), 3963 (2002)
2. Balazevic, I., Allen, C., M. Hospedales, T.: Multi-relational poincare graph embeddings. In: Proceedings of the 33rd NIPS. pp. 4463–4473 (2019)
3. Bonnabel, S.: Stochastic gradient descent on riemannian manifolds. IEEE Transactions on Automatic Control **58**(9), 2217–2229 (2013)
4. Bordes, A., Usunier, N., Garcia-Duran, A., et al: Translating embeddings for modeling multi-relational data. In: Proceedings of the 26th NIPS. pp. 1–9 (2013)
5. Chami, I., Ying, Z., Re, C., et al: Hyperbolic graph convolutional neural networks. In: Proceedings of the 33rd NIPS. pp. 4869–4880 (2019)
6. Kolyvakis, P., Kalousis, A., Kiritsis, D.: Hyperkg: Hyperbolic knowledge graph embeddings for knowledge base completion. CoRR **abs/1908.04895** (2019)
7. Lopez, F., Heinzerling, B., Strube, M.: Fine-grained entity typing in hyperbolic space. In: Proceedings of the 4th RepL4NLP. pp. 169–180 (2019)
8. Sun, Z., Deng, Z., Nie, J., et al: Rotate: Knowledge graph embedding by relational rotation in complex space. In: Proceedings of the 7th ICLR (2019)
9. Ungar, A.A.: Hyperbolic trigonometry and its application in the poincare ball model of hyperbolic geometry. Computers and Mathematics with Applications **41**(1), 135–147 (2001)
10. Yang, B., Yih, W.t., He, X., et al: Embedding entities and relations for learning and inference in knowledge bases. In: Proceedings of the 3rd ICLR (2015)