Constrain the Relational Dimensions for Knowledge Graph Embeddings

Xincan Feng, Hiroyuki Shindo, Yuchang Cheng,
Taro Watanabe, Nobuhiro Yugami

1 Abstract

Knowledge graph embedding (KGE) models have been a popular approach to do link prediction task for knowledge graph completion (KGC). KGE models first choose a representation method, then employ a transformation function to map nodes via edges into a corresponding vector space in order to measure the likelihood of the links. While mapping the individual nodes, the structure of the subgraphs is also transformed. In order to make the KGE models more expressive, researchers are using more and more sophisticated representations, spaces and transformation functions.

However, as those become more complicated, it becomes harder for researchers to distinguish the redundancy of their models. Our research proposed a dimension restriction method to reduce the redundancy in Complex KGE models. To our knowledge, no investigation has explicitly focused on the relational nature in knowlege graphs in the aspect of relational dimensions. Our research has verified and demonstrated the relational nature of dimensions in one of the KGE models: the $5^{\bigstar}\mathrm{E}$ model (Nayyeri et al. 2021). Specifically, we proposed conjugate method to add restriction in the dimensions of parameters matrix, so that to reduce the calculation and learning cost while receiving comparable performance.

2 Introduction

Knowledge graphs (KGs) with their graph-based knowledge representation in the form of (head, relation, tail) triples, have become a leading technology of recent years in AI-based tasks including question answering, data integration, and recommender systems (Ji et al. 2020). However, KGs are incomplete, and the application consuming them is affected by this problem. Knowledge graph embedding (KGE) is a prominent approach used for knowledge graph completion (KGC) by predicting missing links.

Every KGE model defines a specific represention, and uses a transformation function to map entities (nodes) of the KG through relations into a corresponding vector space, and then calculate the plausibility of triples via a corresponding score function. Thus, there are three main directions in improving the KGE models: representations, vector spaces, and transformation functions.

We can generally conclude that:

(i) Compared with real number representation, the composition of complex number can handle a large variety of binary relations, among them symmetric and antisymmetric relations (Trouillon et al. 2016).

(ii) Compared with Euclidean space, the Hyperbolic space models can save more structures using fixed or trainable curvatures in different positions for hierarchical relations (Chami et al. 2020).

(iii) Compared with naive addition or multiplication transformation functions, more specific designed addition, multiplication and their composition, including projective geometric functions, support multiple simultaneous transformations, and are more capable to represent multiple structures in the multi-relational knowledge graphs.

The transformation functions distinguish the extent to which a KGE model is able to learn complicated motifs and patterns formed by combinations of the nodes and edges. While all transformation functions first rest on the representation methods in different spaces. Dimensions, which are the base of the representation methods and vector spaces, are the fundamental base of the transformation functions' capabilities.

Our intuitions are that:
(i) The increased performance from real number representation to complex number representation, which enables the imaginary and real part dimensions to be **relational**, indicates the **hidden relations** between parameter dimensions.
(ii) The increased performance from Euclidean space to Hyperbolic space, which enables the distances or angles of **different positions** to vary in different degrees, indicates **different positions** in the parameter dimensions have different degrees of freedom. Or rather, the fixed or trainable curvatures in Hyperbolic embeddings can learn uneven information to affect the link plausibility result. The uneven parameter dimensions is capable to do that too.

The **"degree of freedom" or call it "restriction" methods on dimensions**, are similar to **"attention" or "weight" methods on parameters** (Vaswani et al. 2017). They all explore the **"relationship problem"** among the feature parameters of the data.
Nonetheless, profoundly speaking, the "restriction" on dimensions wins in that it goes deeper to the root of the "relationship problem", thus is possible to be more effective by reducing the calculation and learning cost.
Concretely speaking, the transformation functions degined in all kinds of representation methods and space, could be **redundant**.

3 Related work

KGE models that are classified according to their representation methods:
Real embeddings Translation approaches include TransE (Bordes et al. 2013) and its variants (Ji et al. 2015, Lin et al. 2015). Although these models are fairly simple and have few parameters, they fail to encode important logical properties (e.g., translations can't encode symmetry).
Complex embeddings ComplEx (Trouillon et al. 2016) gives a clear comparison with respect to existing approaches using only real numbers by presenting an equivalent reformulation of their model that involves only real embeddings. RotatE (Sun et al. 2019) does require additional optimization component to effectively draw negative samples for training. $5^{\bigstar}\mathrm{E}$ (Nayyeri et al. 2021) uses Möbius transformation as the transformation function, which is the composition of four functions that can represent five transformations. However, the parameters usage and the cooperation among parameters in its theoretical analysis are not even. The complex parameters a, b, c, d are used twice, once, four times and once respectively. c cooperates the most with other parameters, whereas b cooperates the least.

KGE models that are classified according to their vector spaces:
Euclidean embeddings Tensor factorization methods such as RESCAL (Nickel et al. 2011) and DistMult (Yang et al. 2015) are designed based on element-wise multiplication of transformed head and tail. In this case, the plausibility of triples is measured based on the angle of transformed head and tail.
Hyperbolic embeddings MuRP (Balazevic et al. 2019) minimizes hyperbolic distances between a re-scaled version of the head entity embedding and a translation of the tail entity embedding. It uses hyperbolic embeddings with fewer dimensions than its Euclidean analogues. ATTH (Chami et al. 2020) leverages trainable hyperbolic curvatures per relationship to simultaneously capture logical patterns and hierarchies. How-

ever, it also needs optimization methods in tangent space (i.e., Euclidean).

4 Method

Our hypothesis is that: The transformation functions of Complex and Hyperbolic embeddings are possibly to be improved in efficiency by dimension restriction methods.
Their performance has already shown improvement compared to Real and Euclidean embeddings by setting relational restrictions in specific positions of the transformation function parameters. Our goal is to find out if there exists redundancy in the transformation function capabilities.

The method that we have varified validity is conjugate method.

5 Experiments and discussion

5.1 Experimental Setup

We followed the best practices of evaluations for KGE models. we consider the most-used metrics Mean Reciprocal Rank (MRR) and Hits@n (n = 1, 3, 10).

We evaluated our method on two widely used benchmark datasets namely FB15k-237 (Toutanova and Chen 2015), and WN18RR (Dettmers et al. 2018). The FB15k-237 and WN18RR datasets both include several relational patterns such as composition (e.g. awardnominee/.../nominatedfor), symmetry (e.g. derivationally_related_form in WN18RR), and anti-symmetry (e.g. has_part in WN18RR). The WN18RR dataset includes hierarchical relations such as hypernym and has_part, which are typical examples for shaping a path structure, and relations such as also_see, similar_to, which are candidates for loop structures.

We compared the results on one of the best performing models $5^\star\mathrm{E}$ (Nayyeri et al. 2021).

Our method is implemented in Pytorch[1] and the code[2] is available online.

Concretely, we restricted the parameters to conjugate of their symmetric counterparts respecting to the main diagonal.

5.2 Results

The parameters size and calculation time were reduced to 75%, while the performances are comparable.

FB15K-237				
Models	MRR	H@1	H@3	H@10
$5^\star\mathrm{E}$	0.3358	0.2462	0.3671	0.5172
$5^\star\mathrm{E}_{\mathrm{conj}}$	0.3351	0.2467	0.3663	0.5132

WN18RR				
Models	MRR	H@1	H@3	H@10
$5^\star\mathrm{E}$	0.4405	0.4123	0.4522	0.4956
$5^\star\mathrm{E}_{\mathrm{conj}}$	0.4386	0.4123	0.4496	0.4890

5.3 Future Work

We will test the stability of experiments results. Moreover, we are going to do more experiments and try out more methods on other KGE models, complex and hyperbolic models especially.

6 References

Nayyeri M.; Vahdati S.; Aykul C.; and Lehmann J. 5★ Knowledge Graph Embeddings with Projective Transformations. In *AAAI*. 2021

Ji S.; Pan S.; Cambria E.; Marttinen P.; and Yu P. S. A survey on knowledge graphs: Representation, acquisition and applications. In *arXiv*:2002.00388. 2020

[1] https://pytorch.org/
[2] https://github.com/FengXincan/dimension

Trouillon T.; Welbl J.; Riedel S.; Gaussier É.; and Bouchard G. Complex embeddings for simple link prediction. In *International Conference on Machine Learning*, 2071-2080. 2016

Chami I.; Wolf A.; Juan D.-C.; Sala F.; Ravi S.; and Ré C. Low-Dimensional Hyperbolic Knowledge Graph Embeddings. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 6901-6914. 2020

Vaswani A.; Shazeer N.; Parmar N.; Uszkoreit J.; Jones L.; Gomez A. N.; Kaiser L.; Polosukhin I. Attention Is All You Need. In *arXiv*:1706.03762v5. 2017

Sun Z.-Q.; Deng Z.-H.; Nie J.-Y.; and Tang J. Rotate: Knowledge graph embedding by relational rotation in complex space. In *International Conference on Learning Representations*. 2019

Bordes A.; Usunier N.; Garcia-Duran A.; Weston J.; and Yakhnenko O. Translating embeddings for modeling multi-relational data. In *Advances in neural information processing systems*, 2787-2795. 2013

Ji G.; He S.; Xu L.; Liu K.; and Zhao J. Knowledge graph embedding via dynamic mapping matrix. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 687-696. 2015

Lin Y.; Liu Z.; Sun M.; Liu Y.; and Zhu X. Learning entity and relation embeddings for knowledge graph completion. In *Twenty-ninth AAAI conference on artificial intelligence*. 2015

Nickel M.; Tresp V.; and Kriegel H.-P. A three-way model for collective learning on multi-relational data. In *International Conference on Machine Learning*, pages 809-816. Omnipress. 2011

Yang B.; Yih W.-t.; He X.; Gao J.; and Deng L. Embedding entities and relations for learning and inference in knowledge bases. In *Conference on Learning Representations (ICLR)*. 2015

Balazevic I.; Allen C.; and Hospedales T. Multi-relational Poincaré graph embeddings. In *Advances in Neural Information Processing Systems*, 4465-4475. 2019

Toutanova K.; and Chen D. Observed versus latent features for knowledge base and text inference. In *Proceedings of the 3rd Workshop on Continuous Vector Space Models and their Compositionality*, 57-66. 2015

Dettmers T.; Minervini P.; Stenetorp P.; and Riedel S. Convolutional 2d knowledge graph embeddings. In *Thirty-Second AAAI Conference*. 2018