Towards Understanding the Geometry of Knowledge Graph Embeddings

Chandrahas, AdityaSharma, Partha Talukdar

Xiaofan Yan

Outline

- Abstract
- Introduction
- Overview of KG Embedding Methods
- Metrics
- Experimental Setup
- Result and Analysis

Abstract

- Geometric understanding of KG embeddings
- and find difference between the geometry by different classes of KG embeddings methods.

Introdction

- Embeddings for KG
  - Geometry of KG embeddings:
    - Length
    - Conicity
    - Study the effects of **model type** and **training hyperparameters** on the geometry of KG embeddings and correlate geometry with performance
  - Contributions
    - Initiate a study to analyze the geometry of KG embeddings
    - Discover insights about geometry of KG embeddings
    - Relationship between geometry attributes

Related Work

- A recent work (MimnoandThompson,2017) is an exception to this which addresses this problem in the context of word vectors.
- Effect of the number of negative samples in KG embedding performance
- KG embedding methods
  - **Additive**
  - **Multiplicative**
  - The entity and relation vectors interact via a neural network

Overview of KG Embedding Methods

- Additive methods (learn embeddings by modeling relations as translation vectors from one entity to another, which results in vectors interacting via the addition operation during training)
  - TransE, TransR and STransE
- Multiplicative methods(quantify the likelihood of a triple belonging to the KG through a multiplicative score function)
  - DistMult, HolE and ComplEx

Type	Model	Score Function $\sigma(h, r, t)$
Additive	TransE (Bordes et al., 2013)	$-\|\mathbf{h}+\mathbf{r}-\mathbf{t}\|_1$
	TransR (Lin et al., 2015)	$-\|M_r\mathbf{h}+\mathbf{r}-M_r\mathbf{t}\|_1$
	STransE (Nguyen et al., 2016)	$-\left\|M_r^1\mathbf{h}+\mathbf{r}-M_r^2\mathbf{t}\right\|_1$
Multiplicative	DistMult (Yang et al., 2014)	$\mathbf{r}^\top(\mathbf{h}\odot\mathbf{t})$
	HolE (Nickel et al., 2016)	$\mathbf{r}^\top(\mathbf{h}\star\mathbf{t})$
	ComplEx (Trouillon et al., 2016)	$\mathbf{Re}(\mathbf{r}^\top(\mathbf{h}\odot\bar{\mathbf{t}}))$

Table 1: Summary of various Knowledge Graph (KG) embedding methods used in the paper. Please see Section 3 for more details.

Metrics

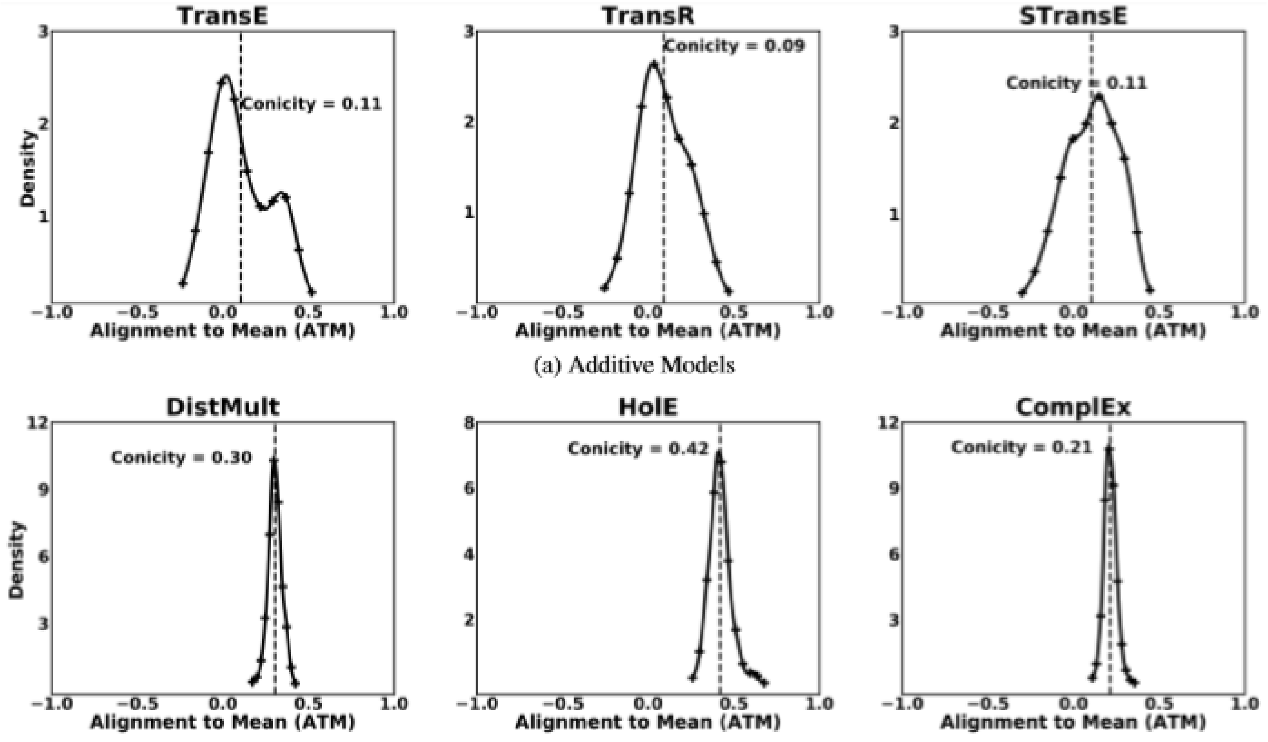
- Alignment to mean (ATM): a vector v belonging to a set of vectors V, as the **cosine similarity** between v and the mean of all vectors in V.
- Conicity: of a set V as the **mean ATM** of all vectors in V.
- vector spread (VS): **variance** of ATM across all vectors in V.
- Average vector length (AVL): the **length** of a vector v as $\boldsymbol{L_2}$-**norm** of the vector $||v_2||$ for the set of the vectors V.

Experimental Setup

- Datasets: Freebase & WorNet (called FC15k & WN18)
- Hyperparameters
- Frequency Bins

Results and Analysis

- Effect of Model Type on Geometry
  - Additive:
    Low conicity and high vector spread.
  - Multiplicative:
    High conicity and low vector spread.

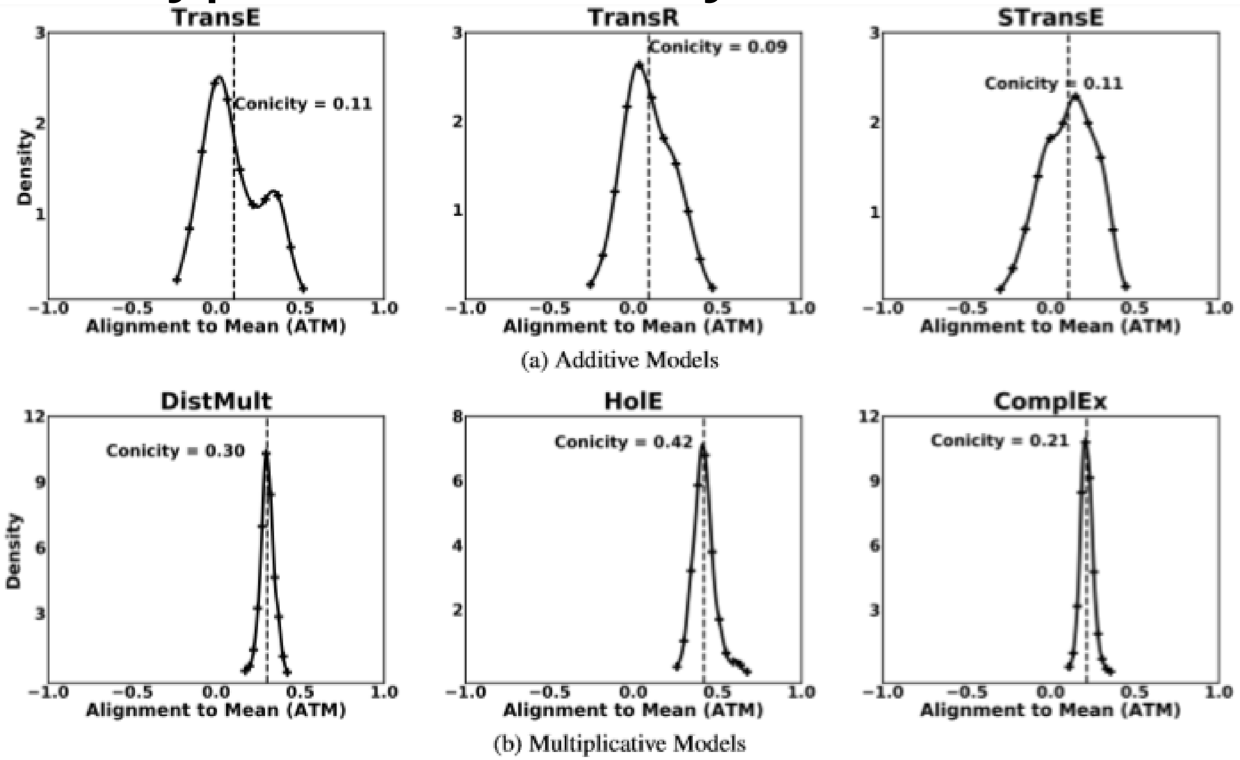


(a) Additive Models
(b) Multiplicative Models

- Alignment to Mean (ATM) vs Density plots for entity embeddings. For each method, a plot averaged across entity frequency bins is shown. We conclude that entity embeddings from additive models tend to have **low (positive as well as negative) ATM** and thereby **low Conicity** and **high vector spread**.

Results and Analysis

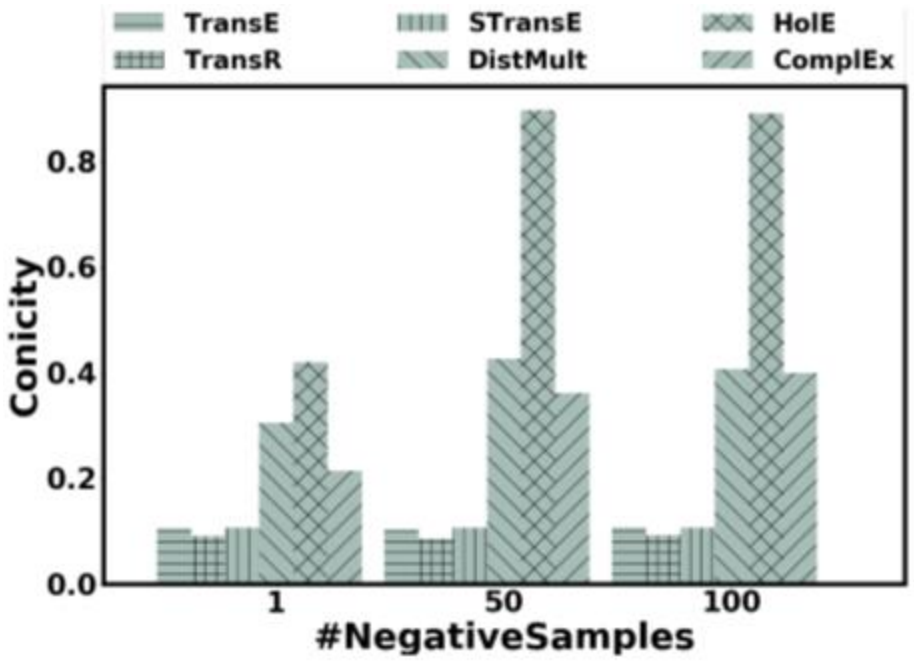
- Effect of Model Type on Geometry

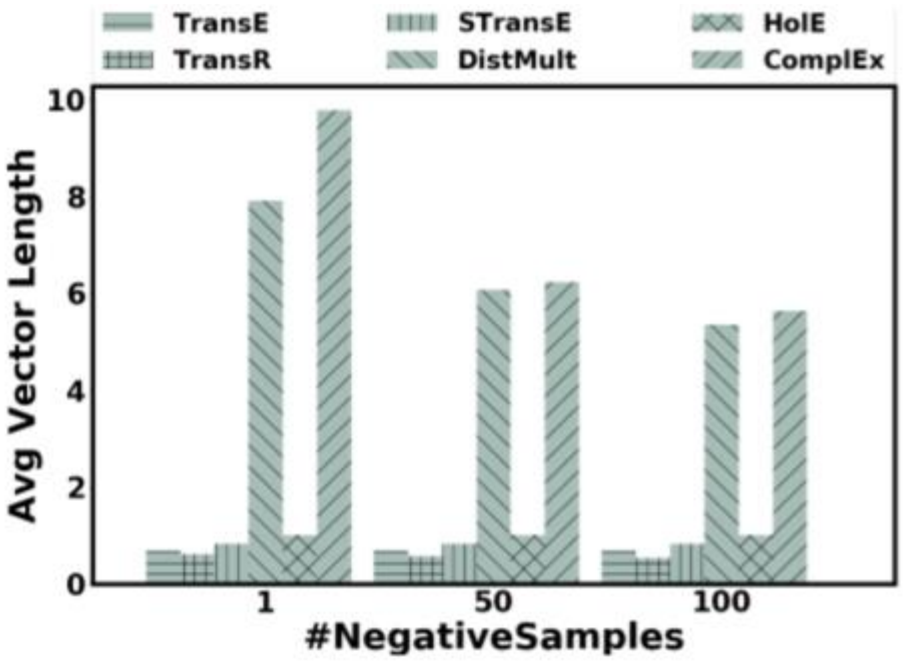


- Alignment to Mean(ATM) vs Density plots for relation embeddings. For each method, a plot averaged across entity frequency bins is shown.

Results and Analysis

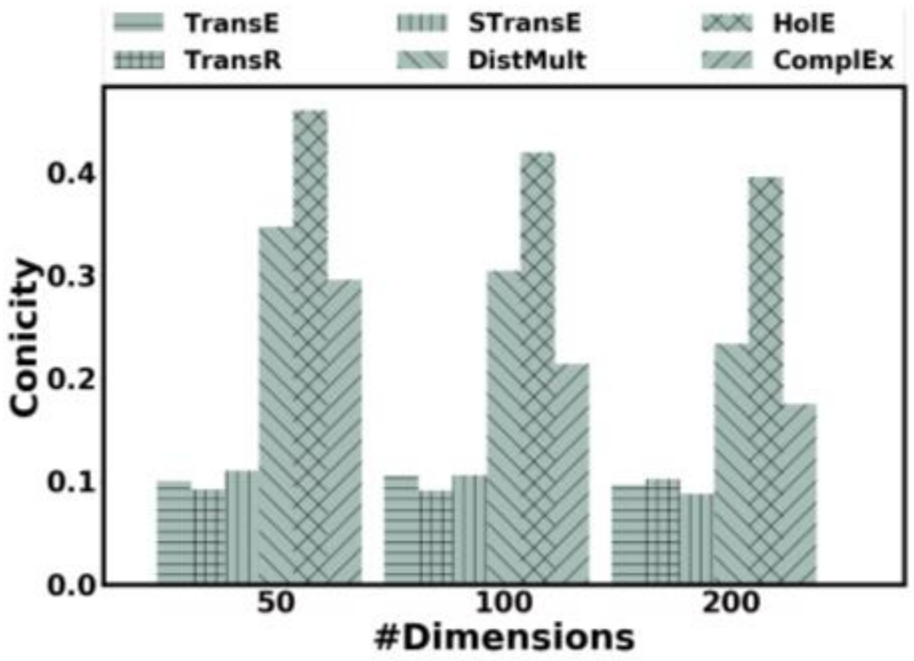
- Effect of Number of Negative Samples on Geometry
  - Additive: Conicity and average length **are invariant to changes** in #NegativeSamples .
  - Multiplicative: Conicity increases while average vector length decrease with increasing #NegativeSamples for entities. Conicity decreases, while average vector length remains constant (except HolE) for relations.

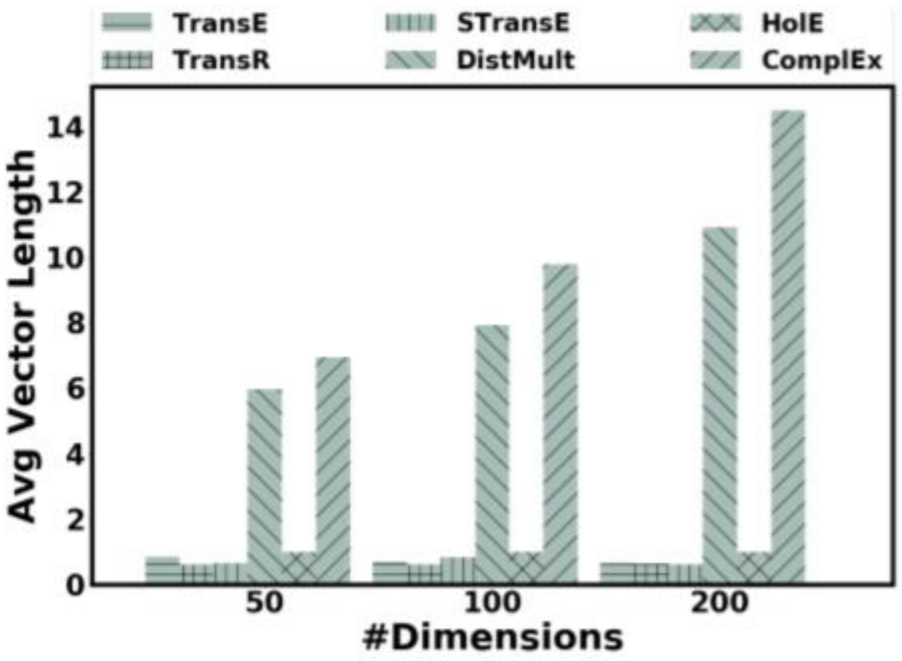




Results and Analysis

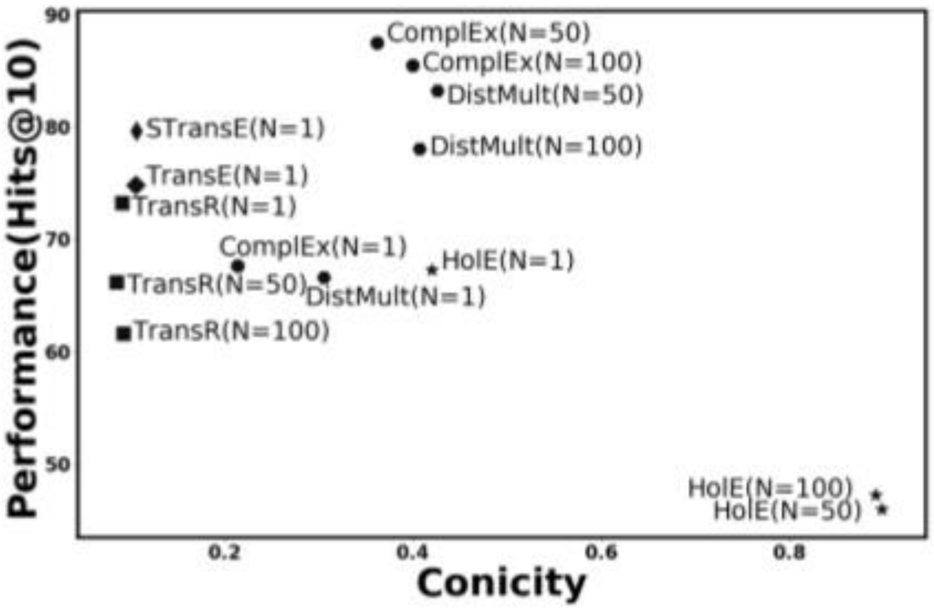
- Effect of Vector Dimension on Geometry
  - Additive: Conicity and average length **are invariant to changes** in dimension.
  - Multiplicative: Conicity decreases for both entities and relations with increasing dimension. Average vector length increases for both entities and relations, except for HolE entities.

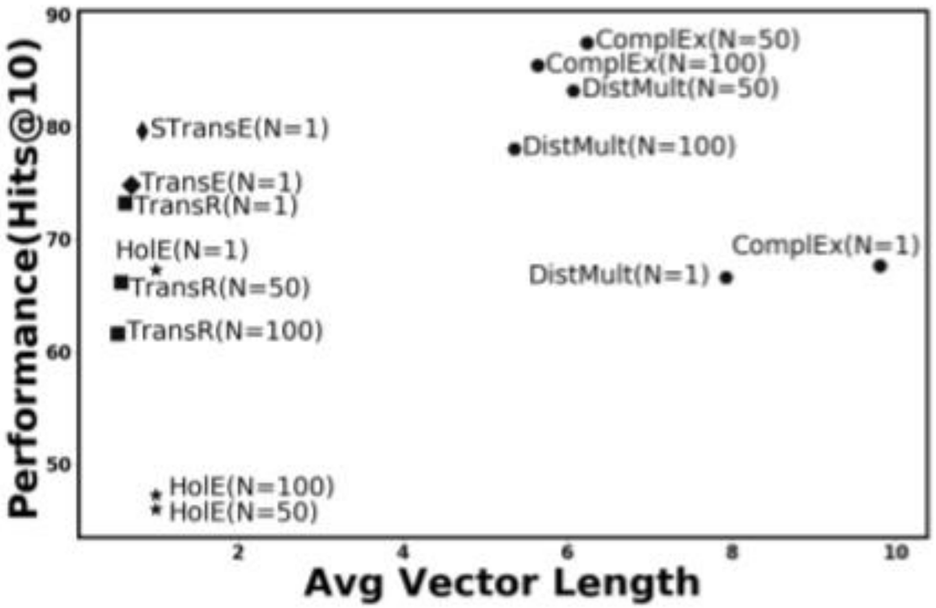




Results and Analysis

- Relating Geometry to Performance
  - Additive: Neither entities nor relations exhibit correlation between geometry and performance.
  - Multiplicative: Keeping negative samples fixed, lower conicity or higher average vector length for entities leads to improved performance. No relationship for relations.





Conclusion

- Initiated a systematic study into the important but unexplored problem of **analyzing geometry of various Knowledge Graph (KG) embedding methods**.