



Complex Embeddings for Simple Link Prediction &

2024-07-30

Presenter : Sooho Moon
DMAIS

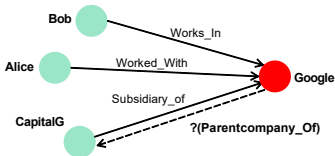
Théo Trouillon, Johannes Welbl, Sebastian Riedel,
Eric Gaussier, Guillaume Bouchard

ICML 2016

- **Main interest**
- **Problem with previous models**
- **Derivation of ComplEx**
- **Architecture**
- **Experiments**
- **Conclusion**

- 지식 그래프에서 **link prediction**을 수행하고자 한다.

- 지식 그래프 : 일반 방향 그래프의 특수한 형태, 노드와(entity), 간선(relation)으로 이루어짐

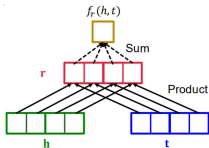


- **Link prediction** : Entity와 relation 사이의 관계를 knowledge base(KB)에서 추론하는 task이며 entity와 relation의 임베딩이 주요 관심사임

Problem with previous models

- Relation을 포괄적으로 학습하지 못한다.

- Relation의 종류는 다양함(symmetry/antisymmetry, inverse, composition)
- 성능이 가장 뛰어났던 모델 중 하나인 DistMult는 antisymmetry를 학습하지 못함

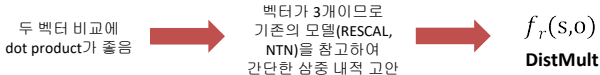


: 삼중 점곱을 수행하는 DistMult

Head entity와 tail entity의 순서가 바뀌어도 결과가 같음
 $\langle h, r, t \rangle = \langle t, r, h \rangle$

■ 일반적인 모델링 방법

- Triple 데이터 (subject, relation, object)에 대해 연관도(점수)를 계산하기 위해 scoring function을 구성함



■ ComplEx의 방법

- Score function을 근사하기 위해 factorize하여 점곱으로 score function을 구성해야함을 보임

Derivation of ComplEx

- 두 entity($|E| = n$)간의 관계를 sign으로 나타내는 부분적으로 관측된 정방 sign matrix Y

$$Y = \begin{bmatrix} -1 & 0 & -1 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix} \quad Y \in \mathbb{R}^{n \times n}$$

연결성 : -1(없음), 0(모름), 1(있음)

- 모델이 예측하는 정방 score matrix는 X (특정 relation에 대한 행렬)로 나타낼 수 있음

$$X = \begin{bmatrix} -2 & 2 & -7 \\ 3 & -1 & 2 \\ 7 & 9 & -10 \end{bmatrix} \quad \text{행(subject)과 열(object)에 대응하는 점수} \quad P(Y_{so} = 1) = \sigma(X_{so})$$

Derivation of ComplEx

- 우리가 원하는 것이 X 이므로 고윳값 분해로 **factorize**함

$$X = EWE^{-1}$$

- 하지만 역행렬을 계산하기가 힘들기 때문에 X 를 대칭 행렬이라고 가정하면 2가지 이점이 있음

1. E 는 **orthogonal**
2. $E^{-1} = E^T$

- 계산은 매우 빨라졌지만 가정 때문에 X 는 **antisymmetric relation**을 표현하지 못함

$$X = \begin{bmatrix} -2 & 0 & 7 \\ 0 & 0 & 2 \\ 7 & 2 & 0 \end{bmatrix}$$

$$(X_{so} =)X_{3,2} = X_{2,3}$$

Derivation of ComplEx

- Antisymmetry를 포함하기 위해 복소 공간으로 이동하고 X 를 normal 한 복소 행렬이라고 가정하여 유니터리 대각화가 가능하게 함

$$X = \text{Re}(EW\bar{E}^T) \quad X, E, W \in \mathbb{C}^{n \times n}$$

- Score값은 실수이므로 복소수의 실수값만 취함
- 결론적으로 새로 정의된 X 는 실행렬이므로 normal일 필요가 없기 때문에 모든 score matrix에 대해 일반화가 가능함

▪ Low-Rank Decomposition

- E 가 정방 행렬이라는 제약이 있음(관측되지 않은 link에 대해서도 일반화 해야함)
- Y 가 low sign-rank라고 가정 $\text{rank}_{\pm}(Y) = \min_{A \in \mathbb{R}^{m \times n}} \{\text{rank}(A) | \text{sign}(A) = Y\}$
- 이 가정을 통해 다음의 결과를 얻음 : $\text{rank}(X) \leq 2\text{rank}_{\pm}(Y)$
- 주로 sign-rank가 매우 작음을 감안하면(e.g., I_n 의 sign-rank는 반드시 3) 아래의 변형이 가능함

$$\text{rank}(EW\bar{E}^T) = K \ll n \quad \longrightarrow \quad E \in \mathbb{C}^{n \times K}, W \in \mathbb{C}^{K \times K}$$

- 하나의 데이터(s, r, o)에 대한 score 계산은 아래와 같음

$$X_{so} = \text{Re}(e_s^T W e_o)$$

$$W \in \mathbb{C}^{K \times K}$$

$$e_s, e_o \in \mathbb{C}^K$$

- 더 general하게 다음과 같이 scoring function을 정의할 수 있음

$$P(Y_{rso} = 1) = \sigma(\phi(r, s, o; \Theta))$$

$$e_s, w_r, e_o \in \Theta$$

$$\min_{\Theta} \sum_{r(s,o) \in \Omega} \log(1 + \exp(-Y_{rso} \phi(s, r, o; \Theta))) + \lambda \|\Theta\|_2^2$$

- Scoring function의 구체적인 계산과 이를 통해 알 수 있는 사실

$$\begin{aligned}\phi(r, s, o; \Theta) &= \text{Re}(\langle w_r, e_s, \bar{e}_o \rangle) \\ &= \text{Re}\left(\sum_{k=1}^K w_{rk} e_{sk} \bar{e}_{ok}\right) \\ &= \langle \text{Re}(w_r), \text{Re}(e_s), \text{Re}(e_o) \rangle \\ &\quad + \langle \text{Re}(w_r), \text{Im}(e_s), \text{Im}(e_o) \rangle \\ &\quad + \langle \text{Im}(w_r), \text{Re}(e_s), \text{Im}(e_o) \rangle \\ &\quad - \langle \text{Im}(w_r), \text{Im}(e_s), \text{Re}(e_o) \rangle\end{aligned}$$

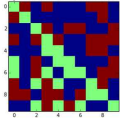
1. Relation 벡터가 순허수면 antisymmetric, 실수면 symmetric을 나타낸다.

2. Symmetric을 나타낼 때 scoring function은 DistMult와 동치다.

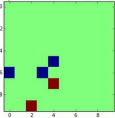
3. 모델의 시공간 복잡도가 KB의 크기에 선형적으로 증가한다.

■ Symmetry, antisymmetry에 대한 성능 확인

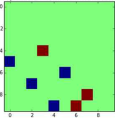
Symmetric relation train slice matrix



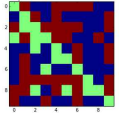
Symmetric relation valid slice matrix



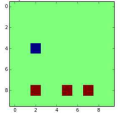
Symmetric relation test slice matrix



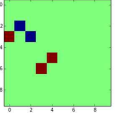
Antisymmetric relation train slice matrix



Antisymmetric relation valid slice matrix



Antisymmetric relation test slice matrix

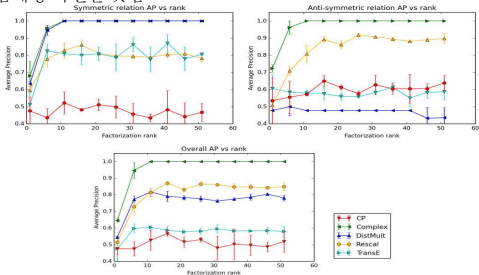


Symmetric relation(상단 3개)과

antisymmetric relation(하단 3개)의 데이터 예시

■ Symmetry, antisymmetry에 대한 성능 확인

- Cross validated Average Precision(area under Precision-Recall curve)중 최댓값
- 가로축은 임베딩 차원을 뜻함



Dataset에 대한 평가

MRR(Mean Reciprocal Rank), **Hits at m**(both are higher, the better)

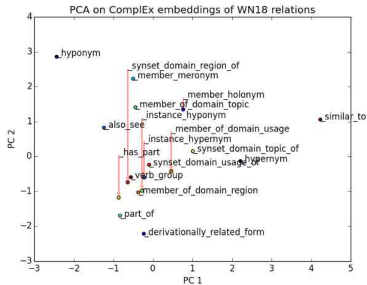
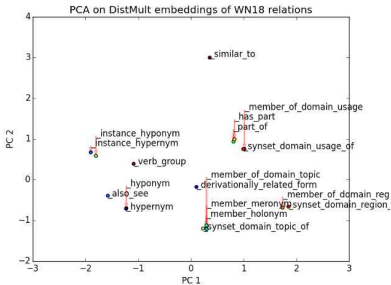
Raw : 초기 데이터

Filter : Rank를 매길 때 실제로 참인 triplet은 metric 계산시 제거
(옳은 예측을 했는데도 penalize 되는 것을 방지)

Model	WN18					FB15K				
	MRR		Hits at			MRR		Hits at		
	Filter	Raw	1	3	10	Filter	Raw	1	3	10
CP	0.075	0.058	0.049	0.080	0.125	0.326	0.152	0.219	0.376	0.532
TransE	0.454	0.335	0.089	0.823	0.934	0.380	0.221	0.231	0.472	0.641
DistMult	0.822	0.532	0.728	0.914	0.936	0.654	0.242	0.546	0.733	0.824
HolE*	0.938	0.616	0.93	0.945	0.949	0.524	0.232	0.402	0.613	0.739
ComplEx	0.941	0.587	0.936	0.945	0.947	0.692	0.242	0.599	0.759	0.840

Experiments

■ 주성분 분석(PCA)



■ Previous problems

Antisymmetric relation 학습에 제약

■ ComplEx

- Antisymmetric relation 학습 가능함
- 복소 공간으로 확장해 다양한 relation을 임베딩할 수 있음
- 수학적 기반으로 모델의 정당성 확보함
- 시공간 복잡도 측면에서 효율적임
- 하지만 composition relation을 학습할 수 없음

Thank You!



Contact: Sooho Moon (Email: moonwalk725@cau.ac.kr)