ニューラルネットワークを用いたグラフ埋め込みによる表現学習

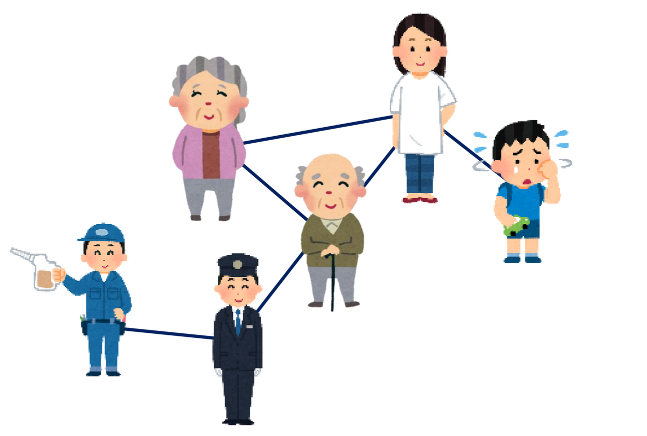
奥野 彰文[1,2]　Kim Geewook[1,2]　下平 英寿[1,2]

[1] 京都大学大学院 情報学研究科 システム科学専攻　[2] 理化学研究所 革新知能統合研究センター（AIP）

1．グラフ埋め込みとは？

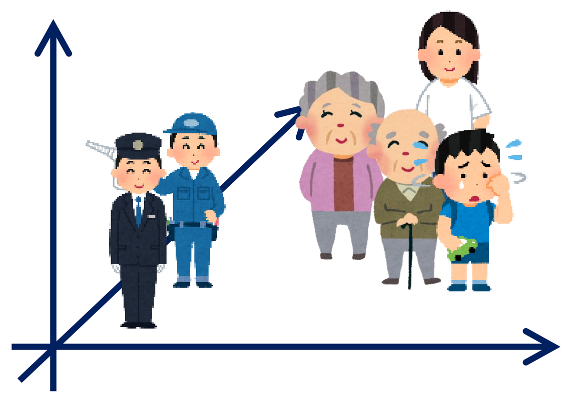
グラフ埋め込みは，**グラフ状の関係性を持ったデータ（ノード）を扱いやすい形式（ベクトル）に変換する方法**です．
あらかじめデータのベクトル表現が得られている場合，**データの次元削減（データ圧縮）ができます**．変換には**ニューラルネットワーク**が使えます．

例1: 友達ネットワークの埋め込み



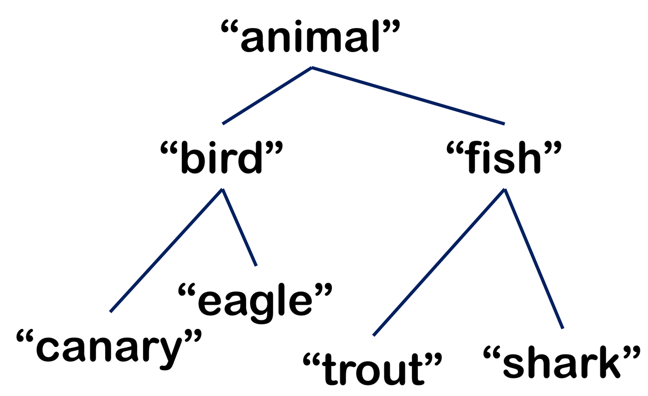
友達なら接続，友達でなければ接続しない．

グラフ埋め込み



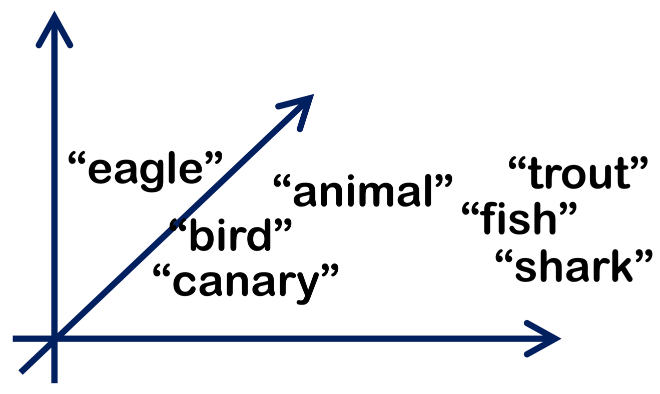
各ノード(人)をユークリッド空間の点として埋め込んだ例．
点と点の距離を測るだけで，友達関係が容易に推測できる．

例2: 英単語 概念辞書の埋め込み (単語埋め込み)



意味の包含関係を表すツリーはグラフとみなせる

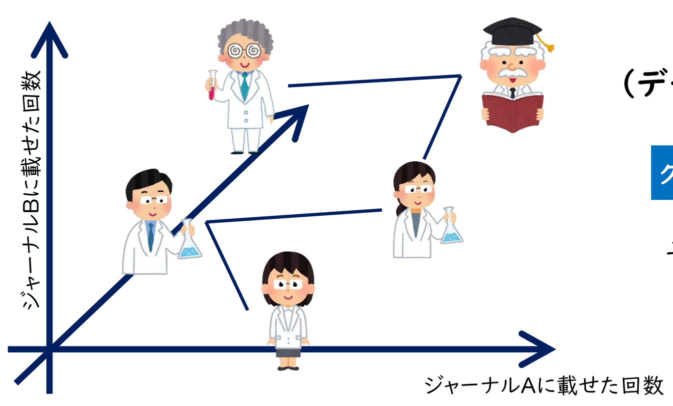
グラフ埋め込み



各ノード(単語)をユークリッド空間の点として埋め込んだ例．
点と点の距離を測るだけで，意味のつながりが推測できる．

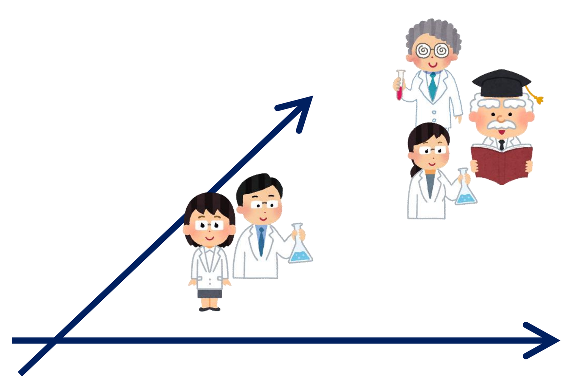
例3: ベクトルが予め得られている 論文共著ネットワークの埋め込み

各ジャーナルに載せた回数を並べて各ノード（研究者）のベクトル表現にする



共著したことがあれば接続，なければ接続しない．

（データの次元削減）
グラフ埋め込み
ニューラルネット



各ノード(研究者)をユークリッド空間の点として埋め込んだ例．
新たに得られたデータベクトルは
（1）グラフ情報を反映し，(2)次元が下がる（データ圧縮）．

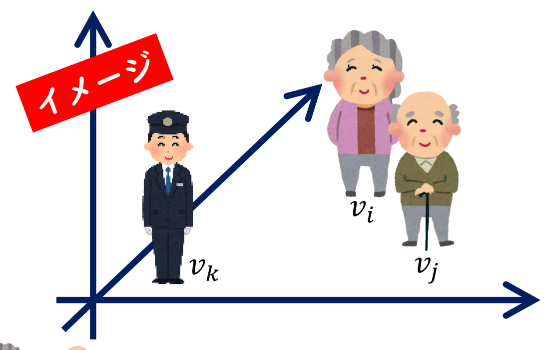
定式化

入力　各ノード v_i のベクトル表現 $\boldsymbol{x}_i \in \mathbb{R}^p$ (未観測なら1-hot)
重み付き隣接行列 $W = (w_{ij})$

$$\ell(\boldsymbol{\theta}) := \sum_{1 \le i < j \le n} \left(w_{ij} - \exp(\langle \boldsymbol{f_\theta}(\boldsymbol{x}_i), \boldsymbol{f_\theta}(\boldsymbol{x}_j)\rangle)\right)^2$$

などを最小化($\boldsymbol{f_\theta}$: ニューラルネットワーク)

出力　各ノード v_i のベクトル表現 $\boldsymbol{y}_i = \boldsymbol{f}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{x}_i) \in \mathbb{R}^K$
(普通は$K \le p$)



v_i と v_j は友達 ($w_{ij}=1$) なので 内積が大きくなる（近くなる）ように，

v_i と v_k は友達ではない ($w_{ik}=0$) ので 内積が小さくなる（遠くなる）ようにする．

やっていることの**まとめ**： と の関係の強さを内積 $\langle \boldsymbol{f_\theta}(\ ), \boldsymbol{f_\theta}(\ )\rangle$ の関数で当てるように埋め込み $\boldsymbol{f_\theta}(\ ) \in \mathbb{R}^K$ を計算

2．我々の貢献

Research Question：$\langle \boldsymbol{f_\theta}(\ ), \boldsymbol{f_\theta}(\ )\rangle$ 以外の類似度を使うとどうなる？

Answer: 埋め込みの次元が高ければ内積を他の正定値類似度に変えても同じ（貢献1）だが，提案した類似度を使えば，より広いクラスの関数を表現できる（貢献2）．

貢献1） ベクトル値ニューラルネットワークの内積 (Inner Product Similarity; IPS) $\langle \boldsymbol{f_\theta}(\boldsymbol{x}_i), \boldsymbol{f_\theta}(\boldsymbol{x}_j)\rangle$ が表現できる関数の広さを調べた．
(Okuno et al., ICML2018 定理 5.1)

主結果：IPS $\langle \boldsymbol{f_\theta}(\ ), \boldsymbol{f_\theta}(\ )\rangle$ は任意の**正定値**類似度 $g(\ ,\ )$ を近似できるが，それ以上は近似できない．

貢献2） 新たな類似度 Shifted IPS (SIPS) $\langle \boldsymbol{f_\theta}(\boldsymbol{x}_i), \boldsymbol{f_\theta}(\boldsymbol{x}_j)\rangle + u_\psi(\boldsymbol{x}_i) + u_\psi(\boldsymbol{x}_j)$ を提案し，IPSより高い表現能力を持つことを示した．
(Okuno and Shimodaira, ICML2018-TADGM WS; Okuno et al., AISTATS2019, to appear)

主結果：提案した**SIPS**は任意の**条件付き正定値**類似度を近似できる！（⇔ **SIPS**はIPSや他の既存法を含む**一般化**）

条件付き正定値類似度（提案したSIPSに対応）

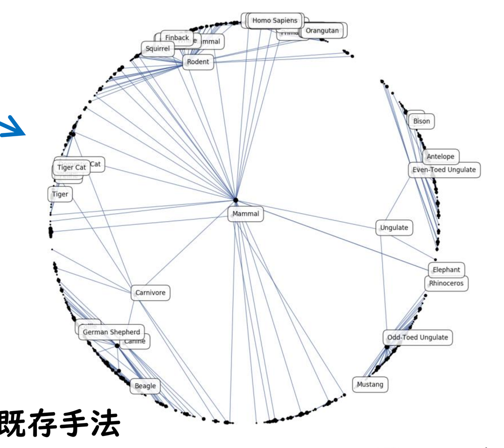
正定値類似度（IPS）

線形カーネル，
多項式カーネル，
ガウスカーネル，…

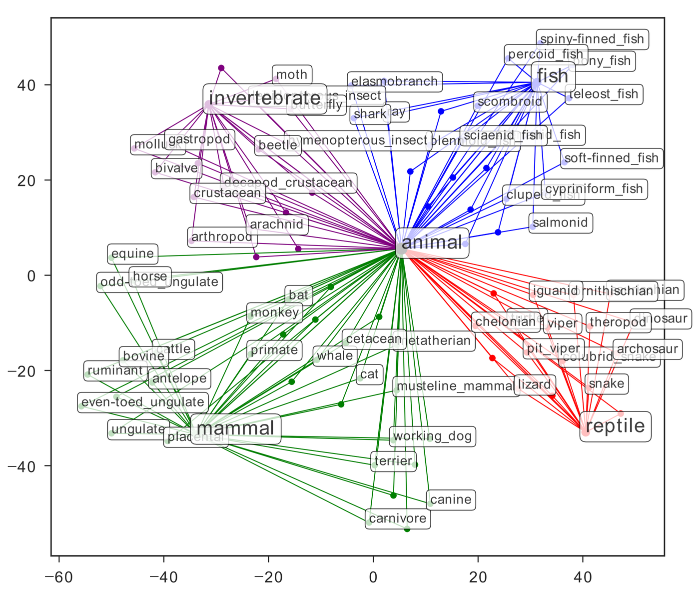
負の二乗距離(NSD)，
ポアンカレ距離(ポアンカレ埋め込み)，
ワッサースタイン距離*[1]，…

*[1]Earth-mover's，またはsliced-Wassersteinに限る

最近，双曲空間でのグラフ埋め込みが注目を集めている：
ポアンカレ埋め込み(Nickel and Kiela, NIPS2017)は
ポアンカレ距離を用いて木(グラフ)を円板上に埋め込む．



既存手法
Nickel and Kiela (2017) 図2(b)より



提案法：Okuno et al. (2019) 図1より

実データでの実験

C-authorship net (Prado et al., 2013) と WordNet (Miller, 1995) の特徴ベクトルを学習し，それぞれでリンク予測，概念辞書の再構成を行い
ROC-AUCスコアを計算した．値は大きいほうが良い．類似度に負の二乗距離（NSD），ポアンカレ距離，IPS，**SIPS（提案）**を用いた．Kは埋め込みの次元．

	Co-authorship network			
	K=2	K=5	K=10	K=20
NSD	**0.822±0.010**	0.866±0.014	0.877±0.012	0.865±0.033
Poincare	0.707±0.021	0.874±0.001	0.882±0.001	0.884±0.001
IPS	0.780±0.005	0.883±0.001	0.896±0.001	0.896±0.001
SIPS (**提案**)	0.781±0.001	**0.885±0.001**	**0.896±0.002**	**0.897±0.001**

	WordNet			
	K=2	K=5	K=10	K=20
NSD	0.792±0.007	0.900±0.001	0.957±0.001	0.984±0.000
Poincare	0.840±0.007	**0.979±0.001**	**0.987±0.000**	0.985±0.000
IPS	0.725±0.006	0.760±0.006	0.769±0.002	0.792±0.002
SIPS (**提案**)	**0.963±0.001**	0.977±0.001	0.983±0.001	**0.987±0.000**

理論

Okuno et al. (AISTATS2019) 定理5.2

gを条件付き正定値カーネル，$f_* \in \mathbb{R}^{K_*}$ をベクトル値連続関数，$S(W_f, K)$を層の数$O(W_f)$で幅が$O(1)$のディープニューラルネットの集合とすると

$$\inf_{\substack{f \in S(W_f, K) \\ u \in S(W_u, 1)}} \sup_{x, x' \in X} \left| g\big(f_*(x), f_*(x')\big) - \big(\langle f(x), f(x')\rangle + u(x) + u(x')\big) \right| = O\left(K^{-\frac{1}{K_*}} + K^{\frac{1}{2}+\frac{2}{p}} W_f^{-\frac{2}{p}} + W_u^{-\frac{2}{p}}\right).$$

参考文献

1. Nickel, Maximillian, and Douwe Kiela. "Poincaré Embeddings for Learning Hierarchical Representations." *Advances in Neural Information Processing Systems (NeurIPS)*. 2017.
2. Okuno, Akifumi, Tetsuya Hada, and Hidetoshi Shimodaira. "A probabilistic framework for multi-view feature learning with many-to-many associations via neural networks," *Proceedings of the 35th International Conference on Machine Learning (ICML)*, Stockholm, Sweden, 2018, PMLR 80:3888-3897.
3. Okuno, Akifumi, and Hidetoshi Shimodaira. "On Representation Power of Neural Network-Based Graph Embedding and Beyond," ICML 2018 workshop on Theoretical Foundations and Applications of Deep Generative Models, *CoRR abs/1805.12332*, 2018
4. Okuno, Akifumi, Geewook Kim, and Hidetoshi Shimodaira. "Graph Embedding with Shifted Inner Product Similarity and Its Improved Approximation Capability," *CoRR abs/1810.03463*, (to appear in AISTATS2019)

Contact info: A. Okuno
http://okuno.co