

Matrix Factorization and Node Embedding

考虑一种这样的相似性度量方式, $\text{similarity}(i,j) = A_{ij}$,

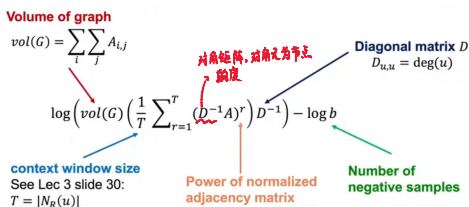
对于节点嵌入来说, 这意味着 $Z^T Z = A$ (A 的矩阵分解)

精确求解通常不可能

因此有以下目标函数: $\min_Z \|A - Z^T Z\|_F$

同理, 对于 DeepWalk, 求节点嵌入的过程等价于进行以下的矩阵分解:

$$D = \log(\text{vol}(G)) \left(\frac{1}{T} \sum_{r=1}^T (D^{-1} A)^r \right) D^{-1} - \log b$$



所以基于随机游走的节点嵌入也可以转化为以下优化问题: $\min_Z \|D - Z^T Z\|_F$

此前涉及的所有节点嵌入方法均无法为训练集中不存在的节点获取嵌入, 此外, 无法捕获局部结构相似性, 无法利用额外的各级别特征。