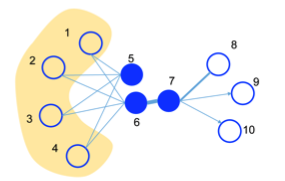
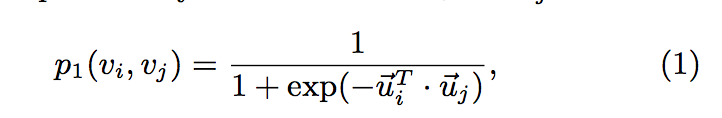
# 基础知识

一阶相似 二阶相似 无向(是特殊的有向)



# 一阶相似度



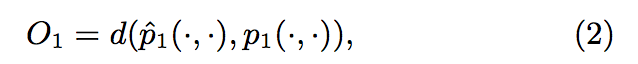
vi vj 的顶点的联合概率的表示

ui uj 是两个顶点的向量表示

p(.,.) 是在V\*V 整个空间的两个边的经验概率定义下面的公式

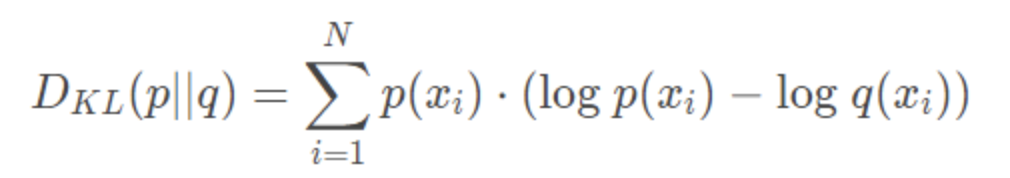
 

我们的目标函数就是

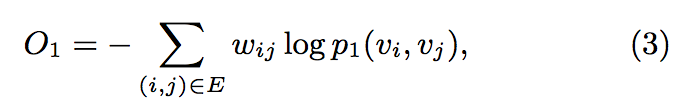


这个是算顶点的联合概率分布和边的概率分布的距离最小

根据KL 散度的距离公式



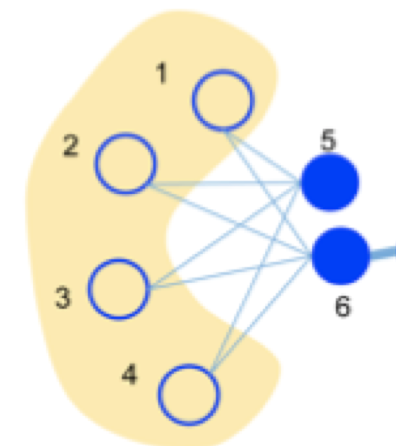
变换目标函数为



这样就能得到ui uj 的向量表示

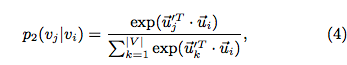
# 二阶相似度

适用于有向和无向图,再给定的网络是有向的(无向就是特殊的有向)



是自己本身的向量表示

是上下文(我理解的就是例如5 的上下文就是1234)



其中|V| 是上下文的数量也就是顶点数

真对于每个顶点vi 定义条件分布概率为

我们的目标函数



由于网络中的顶点的重要性可能是不同的,所以有

可以看一下其他的文献 15



wij 是边I 和j 的权重

di 是顶点I 的 out-degree(出度和入度) 

n(i) 是顶点vi 的外邻集

为了方便计算 把 换成

然后用KL 公式替换D()的函数..就成了下面的函数



最小化目标函数就能得到顶点I 的低维向量