2. 知识图谱表示学习模型

知识表示模型将优化条件概率 $P((h,r,t) \mid \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$ 转化为优化 $P(h \mid (r,t), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$、$P(t \mid (h,r), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$ 及 $P(r \mid (h,t), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$。

对于每一个知识图谱 G 中的实体对 (h,t)，这里定义一个潜在关系向量 $\boldsymbol{r}_{ht}$ 来表达实体向量 $\boldsymbol{h}$ 到实体向量 $\boldsymbol{t}$ 之间的变换与关联，具体形式如下：

$$\boldsymbol{r}_{ht} = \boldsymbol{t} - \boldsymbol{h} \tag{3.53}$$

与此同时，对于知识图谱 $\mathcal{G}$ 中的任意三元组 $(h,r,t) \in \mathcal{T}$，对应存在一个显式的关系 r 来描述 h 与 t 的关系，且这个 r 存在一个显式关系向量 $\boldsymbol{r}$。所以，这里可以将三元组的能量函数定义为

$$f_r(h,t) = b - \|\boldsymbol{r}_{ht} - \boldsymbol{r}\| = b - \|(\boldsymbol{t} - \boldsymbol{h}) - \boldsymbol{r}\| \tag{3.54}$$

其中，b 是一个常数偏移量。基于这个能量函数，这里以 $P(h \mid (r,t), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$ 为例来形式化地给出 $\mathcal{T}$ 中三元组的条件概率：

$$P(h \mid (r,t), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R) = \frac{\exp(f_r(h,t))}{\sum\limits_{h' \in E} \exp(f_r(h',t))} \tag{3.55}$$

类似地，可以定义 $P(t \mid (h,r), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$ 和 $P(r \mid (h,t), \boldsymbol{\theta}_E, \boldsymbol{\theta}_R)$。实际上，无论是出于理念还是落实到具体模型上，这个条件概率所表达的任务和 TransE 是一致的，只是其不再是基于边界值优化而是基于条件概率优化，但本质上没有差别。因此，我们将这个知识图谱表示学习模型命名为 Prob-TransE，寓意概率形式的 TransE。

为了体现联合学习模式可以适应多种知识图谱表示学习模型，这里引入了 TransD[91] 来对知识图谱中的三元组进行编码和嵌入，具体形式如下：

$$\begin{cases} \boldsymbol{r}_{ht} = \boldsymbol{t}_r - \boldsymbol{h}_r \\ \boldsymbol{h}_r = \boldsymbol{M}_{rh}\boldsymbol{h}, \quad \boldsymbol{t}_r = \boldsymbol{M}_{rt}\boldsymbol{t} \\ \boldsymbol{M}_{rh} = \boldsymbol{r}_p \boldsymbol{h}_p^{\top} + \boldsymbol{I}^{k_r \times k_w} \\ \boldsymbol{M}_{rt} = \boldsymbol{r}_p \boldsymbol{t}_p^{\top} + \boldsymbol{I}^{k_r \times k_w} ① \end{cases} \tag{3.56}$$

其中，$\boldsymbol{r}_p$、$\boldsymbol{h}_p$、$\boldsymbol{t}_p$ 都是用来进行映射的工作向量。类似于 Prob-TransE，我们将基于 TransD 进行条件概率优化的知识图谱表示学习模型命名为 Prob-TransD。

① k_r、k_w 分别是关系与实体向量维度。