On Leveraging Variational Graph Embeddings for Open World Compositional Zero-Shot Learning

Anonymous CVPR 2022 submission

Paper ID 7483

A. Supplementary Material: Derivation of ELBO Bound

We aim to learn the free parameters of our model such that the log probability of $\mathcal{G}$ is maximized *i.e.*

$$\begin{aligned} log\Big(p(\mathcal{G})\Big) &= log\Big(\int p(\boldsymbol{Z})p_\theta(\mathcal{G}|\boldsymbol{Z})\, d\boldsymbol{Z}\Big) \\ &= log\Big(\int \frac{q_\phi(\boldsymbol{Z}|\mathcal{G})}{q_\phi(\boldsymbol{Z}|\mathcal{G})} p(\boldsymbol{Z})p_\theta(\mathcal{G}|\boldsymbol{Z})\, d\boldsymbol{Z}\Big) \\ &= log\Big(\mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\{\frac{p(\boldsymbol{Z})p_\theta(\mathcal{G}|\boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\}\Big), \end{aligned} \tag{1}$$

In order to ensure computational tractability, we use Jensen's Inequality [?] to get ELBO bound of Eq. (1). i.e.

$$log\Big(p(\mathcal{G})\Big) \geq \mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\{log\Big(\frac{p(\boldsymbol{Z})p_\theta(\mathcal{G}|\boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big)\Big\} \tag{2}$$

$$\begin{aligned} &= \mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\{log\Big(p_\theta(\mathcal{G}|\boldsymbol{Z})\Big)\Big\} \\ &+ \mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\{log\Big(\frac{p(\boldsymbol{Z})}{q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big)\Big\} \end{aligned} \tag{3}$$

We follow Kipf *et al*. [?] and restrict the decoder $p_\theta(\mathcal{G}|\boldsymbol{Z})$ to reconstruct only edge information from the latent space. The edge information is contained in the adjacency matrix $\boldsymbol{A}$. In other words, we choose the decoder to be an edge decoder *i.e.* $p_\theta(\boldsymbol{A}|\boldsymbol{Z})$.

$$\begin{aligned} log\Big(p(\mathcal{G})\Big) &\geq \mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\{log\Big(p_\theta(\boldsymbol{A}|\boldsymbol{Z})\Big)\Big\} \\ &- D_{KL}\Big(q_\phi(\boldsymbol{Z}|\mathcal{G})||p(\boldsymbol{Z})\Big) \end{aligned} \tag{4}$$

where, D_{KL} denotes the Kullback-Leibler (KL) divergence between the prior and approximate posterior distributions. By using (**??**), (**??**), (**??**) and (**??**), the loss function can be formulated as negative of ELBO bound (4) *i.e.*

$$\begin{aligned} \mathcal{L}_{\text{ELBO}} &= \sum_{i=1}^{N} D_{KL}\Big(\mathcal{N}\Big(\boldsymbol{\mu}_i(\mathcal{G}), \boldsymbol{\sigma}_i^2(\mathcal{G})\Big)\,||\,\mathcal{N}(\boldsymbol{0}, \text{diag}(\boldsymbol{1}))\Big) \\ &- \mathbb{E}_{\boldsymbol{Z}\sim q_\phi(\boldsymbol{Z}|\mathcal{G})}\Big\{log\Big(p_\theta(\boldsymbol{A}|\boldsymbol{Z})\Big)\Big\}. \end{aligned}$$