Densifier

In this section, we briefly introduce *Densifier*, which learns an orthogonal transformation of embedding vectors such that task-relevant information is organized in an ultradense subspace. In our experiments, we apply *Densifier* to transform sentiment information into the ultradense subspace. However, research results reflect that *Densifier* is also applicable to concreteness and frequency, as suggested in Rothe, Ebert, and Schütze (2016).

Given a word w and its d dimensional embedding vector e_w, *Densifier* computes a new d^* dimension vector representation u_w^* for w:

$$u_w^* := P^* Q e_w$$

where $Q \in \mathbb{R}^{d \times d}$ is an orthogonal matrix. $P^* \in R^{d^* \times d}$ is an identity matrix for dimensions $\{1, 2, ... \, d^*\}$ which serves as a selector. Generally the subspace size d^* is much smaller than d, and in extreme case d^* is set to 1 to yield lexicons.

Given positive sentiment word group G_p and negative sentiment word group G_n in a sentiment lexicon $\mathcal{D}$, *Densifier* tries to find an orthogonal matrix Q such that word pairs in the same sentiment group (G_{same}) have small Euclidean distance in the ultradense subspace and *vice versa*. More concretely, the objective is $\forall w, v \in \mathcal{D}, w \neq v$:

$$\underset{Q}{\arg\min} \sum_{(w,v) \in G_{diff}} -\alpha^* \|P^* Q (e_w - e_v)\|_2 \quad + \sum_{(w,v) \in G_{same}} (1 - \alpha^*) \|P^* Q (e_w - e_v)\|_2$$

s.t. Q is an orthogonal matrix. α^* is the hyperparameter balancing relative importances of the two terms. We then minimize the loss by stochastic gradient descent (SGD), where gradient $\nabla_Q(L)$ is computed as:

$$\nabla_Q(L) := \sum_{(w,v) \in G_{diff}} -\alpha^* \nabla_Q \|P^* Q (e_w - e_v)\|_2 \quad + \sum_{(w,v) \in G_{same}} (1 - \alpha^*) \nabla_Q \|P^* Q (e_w - e_v)\|_2$$

Denoting $e := e_w - e_v$, we note that:

$$\begin{aligned}
\nabla_Q \|P^* Q e\|_2 &= \nabla_Q [(P^* Q e)^T (P^* Q e)]^{\frac{1}{2}} \\
&= \nabla_Q [e^T Q^T P^{*T} P^* Q e]^{\frac{1}{2}} \\
&= \frac{1}{2} [e^T Q^T P^{*T} P^* Q e]^{-\frac{1}{2}} \nabla_Q (e^T Q^T P^{*T} P^* Q e) \\
&= [e^T Q^T P^{*T} P^* Q e]^{-\frac{1}{2}} (P^{*T} P^* Q e e^T) \\
&= \frac{P^{*T} P^* Q e e^T}{\|P^* Q e\|_2}
\end{aligned}$$

such that we could compute $\nabla_Q(L)$ accordingly.

We then update Q:

$$Q := Q - \beta \nabla_Q (L)$$

$$U S V^T = \texttt{svd}(Q)$$