

# Building an initial solution

$X$  and  $Z$  be the word embedding matrices in two languages, so that their  $i$ th row  $X_i^*$  and  $Z_i^*$  denote the embeddings of the  $i$ th word in their respective vocabularies.

We aim to build a dictionary between both languages, encoded as a sparse matrix  $D$  where  $D_{ij} = 1$  if the  $j$ th word in the target language is a translation of the  $i$ th word in the source language.

Embedding Normalization:

- Length Normalizing the embeddings.
- Mean centering each dimension.
- Length Normalizing again in order to make all embeddings as unit vector. With unit vectors, dot products are nothing but the cosine similarity.

Fully unsupervised initialization:

- The underlying difficulty of the mapping problem in its unsupervised variant is that the word embedding matrices  $X$  and  $Z$  are unaligned across both axes: neither the  $i$ th vocabulary item  $X_i^*$  and  $Z_i^*$  nor the  $j$ th dimension of the embeddings  $X_j^*$  and  $Z_j^*$  are aligned, so there is no direct correspondence between both languages.
- In order to overcome this challenge we construct 2 alternative representations  $X'$  and  $Z'$  that are aligned along their  $j$ th dimension. These matrices then can be used to align their respective vocabularies.
- The idea is to construct similarity matrices  $MX$  and  $MZ$  from the respective  $X$  and  $Z$  embedding matrices. The advantage of using similarity matrices is that both the axes are words and hence aligning their  $j$ th dimensions becomes easy.