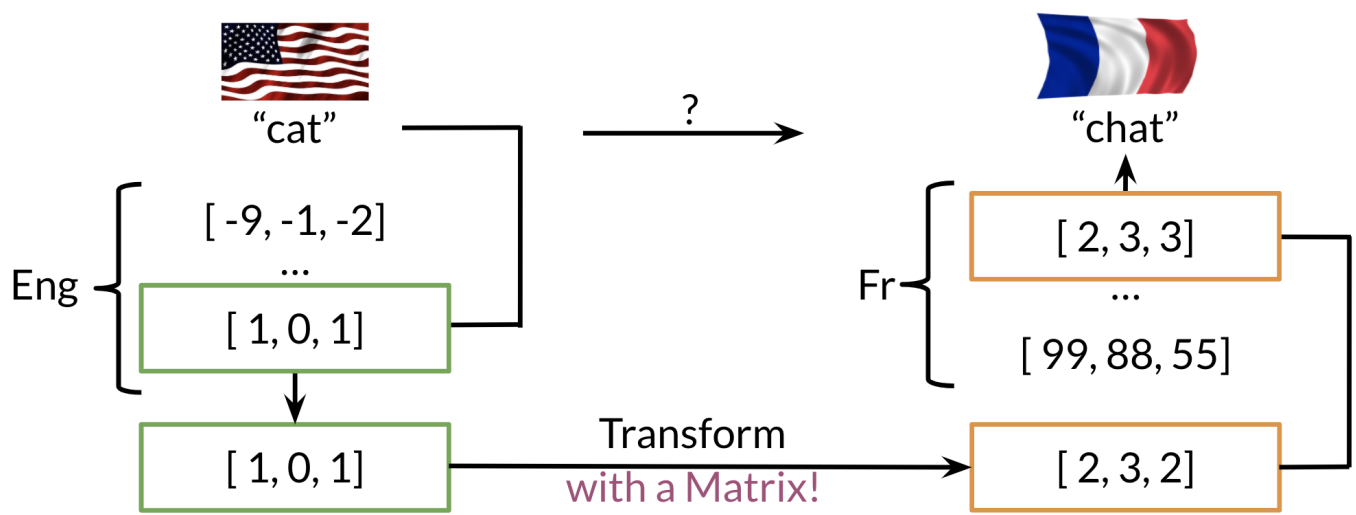




Item Navigation

Transforming word vectors

In the previous week, I showed you how we can plot word vectors. Now, you will see how you can take a word vector and learn a mapping that will allow you to translate words by learning a "transformation matrix". Here is a visualization:



Note that the word "chat" in french means cat. You can learn that by taking the vector corresponding to "cat" in english, multiplying it by a matrix that you learn and then you can use cosine similarity between the output and all the french vectors. You should see that the closest result is the vector which corresponds to "chat".

Here is a visualization of that showing you the aligned vectors:

$$\begin{pmatrix} \text{["cat" vector]} \\ \text{[... vector]} \\ \text{["zebra" vector]} \end{pmatrix} \mathbf{X} \mathbf{R} \approx \mathbf{Y} \begin{pmatrix} \text{["chat" vecteur]} \\ \text{[... vecteur]} \\ \text{["z bresse" vecteur]} \end{pmatrix} \mathbf{Y}$$

subsets of the full vocabulary

Note that X corresponds to the matrix of english word vectors and Y corresponds to the matrix of french word vectors. R is the mapping matrix.

Steps required to learn R :

- Initialize R
- For loop

$$Loss = ||XR - Y||_F$$

$$g = \frac{d}{dR} Loss$$

$$R = R - \alpha * g$$

Here is an example to show you how the frobenius norm works.

$$\begin{aligned} &||\mathbf{XR} - \mathbf{Y}||_F \\ \mathbf{A} &= \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} \\ ||\mathbf{A}_F|| &= \sqrt{2^2 + 2^2 + 2^2 + 2^2} \\ ||\mathbf{A}_F|| &= 4 \\ ||\mathbf{A}||_F &\equiv \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} \end{aligned}$$

In summary you are making use of the following: