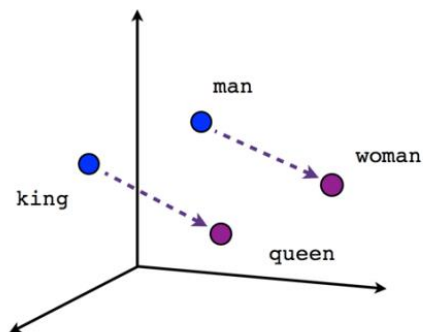
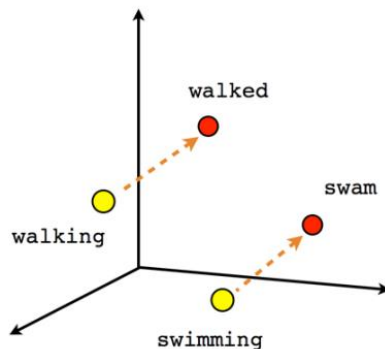


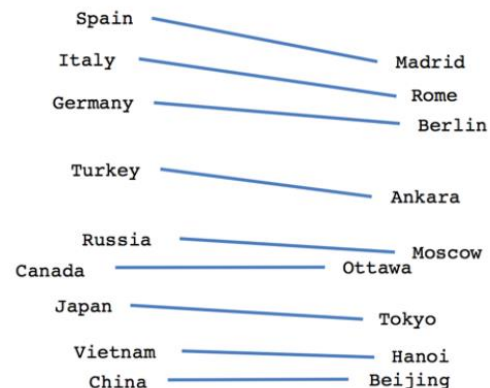
Word2Vec



Male-Female



Verb tense



Country-Capital

$$\text{vector[Queen]} = \text{vector[King]} - \text{vector[Man]} + \text{vector[Woman]}$$

Word2Vec

- Objective: find a 128-dimensional vector $word2vec(w)$ that represents each word w in a “useful” way
 - “Useful” can mean we want to distinguish pairs of words that occur together from pairs of words that don’t by looking at the vector representations of the words
 - “Useful” can mean $word2vec(w)$ can be used to predict whether the word w occurs in positive or negative reviews
 - Even if in the training set, w doesn’t occur in either!

Learning Word2Vec with negative sampling

- Learn to distinguish words that occur together from words that don't occur together:
Maximize the following with w_1 and w_2 occurring together in the text, and w_i' s not occurring with w_1

$$\log \sigma(v_{w_1}^T v_{w_2}) + \sum_i \log \sigma(-v_{w_1}^T v_{w_i'})$$

- Encourage $v_{w_1}^T v_{w_2}$ to be high, encourage $-v_{w_1}^T v_{w_i'}$ to be low
- $v_{w_i'}$ are sampled randomly from the text, with more frequent words sampled more frequently

Using Word2Vec for Sentiment Analysis

- The easiest thing: average the word2vec representations of all the words in the review, classify those
- Standard approach: use Recurrent Neural Networks
 - Coming up!

Word2Vec: unsupervised learning

- Very easy to get large amounts of unlabelled text in order to learn really good Word2Vec representations
- Can then use the word representations to learn from small datasets