**Word Embedding**

Analogies using Word Vectors:

Cosine Similarity:

Embedding Matrix:

: the embedding matrix, shape =

: the one-hot vector of the *i*-th word, shape =

: the embedding vector of the *i*-th word, shape =

**Word2Vec**

Skip-Grams Softmax:

: the context word (input)

: the target word (output)

: the embedding vector of the context word, shape =

: the parameter associated with the target word

Skip-Grams Loss Function:

**Negative Sampling**

|  |  |  |
| --- | --- | --- |
| context | word | target? |
| orange | juice | 1 |
| orange | king | 0 |
| orange | book | 0 |
| orange | the | 0 |
| orange | of | 0 |
| ↑ | ↑ | ↑ |
| ***c*** | ***t*** | ***y*** |

Logistic Regression:

**GloVe (Global Vectors for Word Representation)**

Minimize:

: the number of times of word *i* appearing in context of word *j*

**Sentiment Classification**