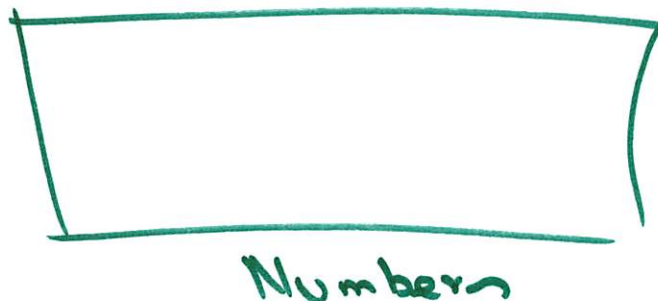
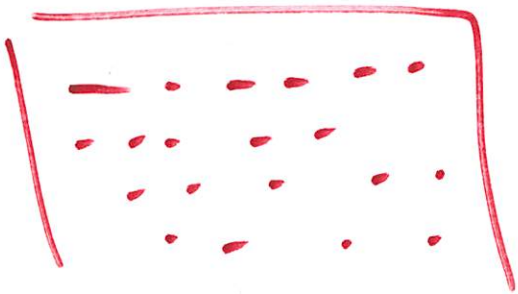


Math

Crop
Brighten
rotation



Math

translation
Sentiment
...

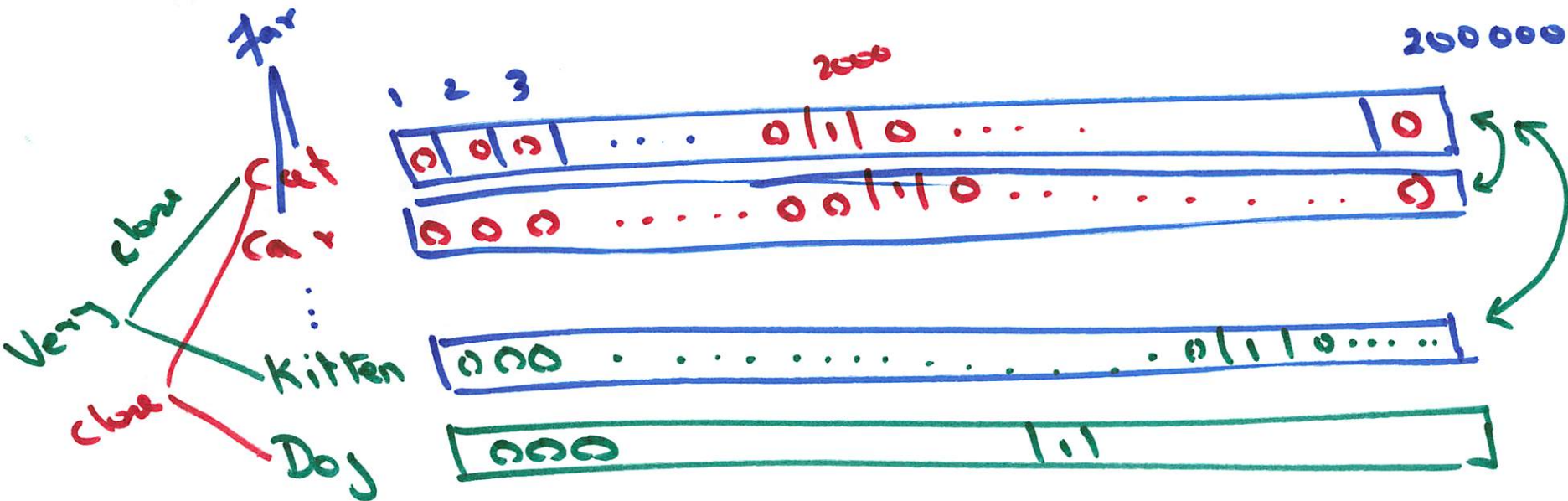
One hot encoding

200,000 words

Eucledian

$$\frac{\Delta y}{\Delta x} \quad \sqrt{\Delta x^2 + \Delta y^2}$$

Cos Similinh



Word Embeddings

- Word Embeddings: An NLP technique of representing words as vectors
- Several methods exist to do this kind of translation. Popular methods include

→ ☹️ One Hot encoding ←

→ • Dimension reduction techniques, like SVD ←

- word2Vec (Neural networks)
- GloVe

words → Vectors

7 million articles
5 billion words
→ 250 million sent

- I went to the cinema on Sunday
- I went to the beach on Monday

- My favorite pet is a cat
- I like walking my pet dog

- I bought a red Honda
- I bought a red car
- I bought a blue Toyota

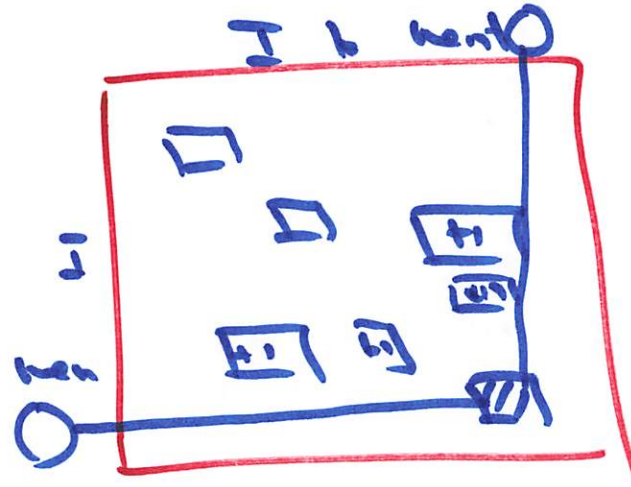
Dim Reduction

CoOccurance Matrix

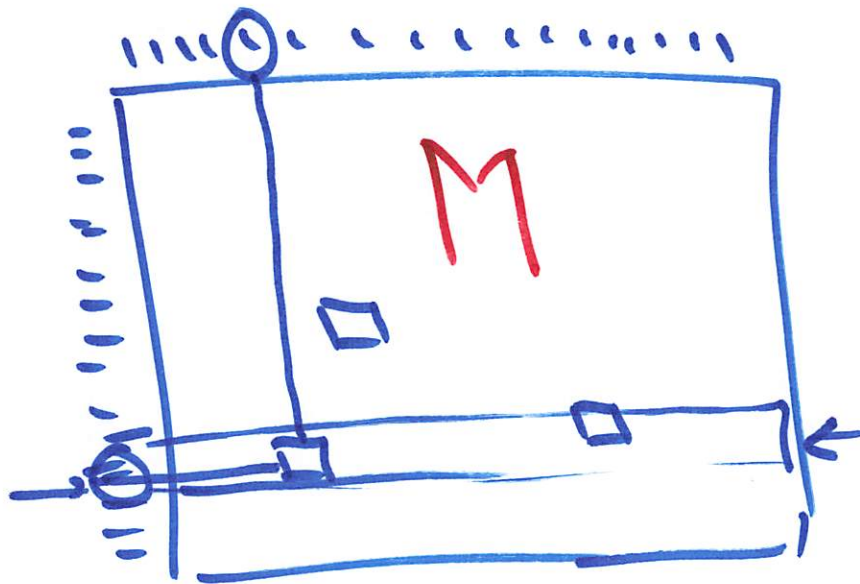
7 million

5 billion words

250 million sent.

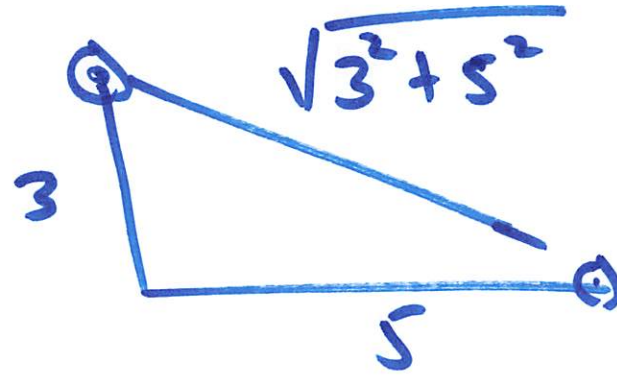


I went to the
beach on Sunday.

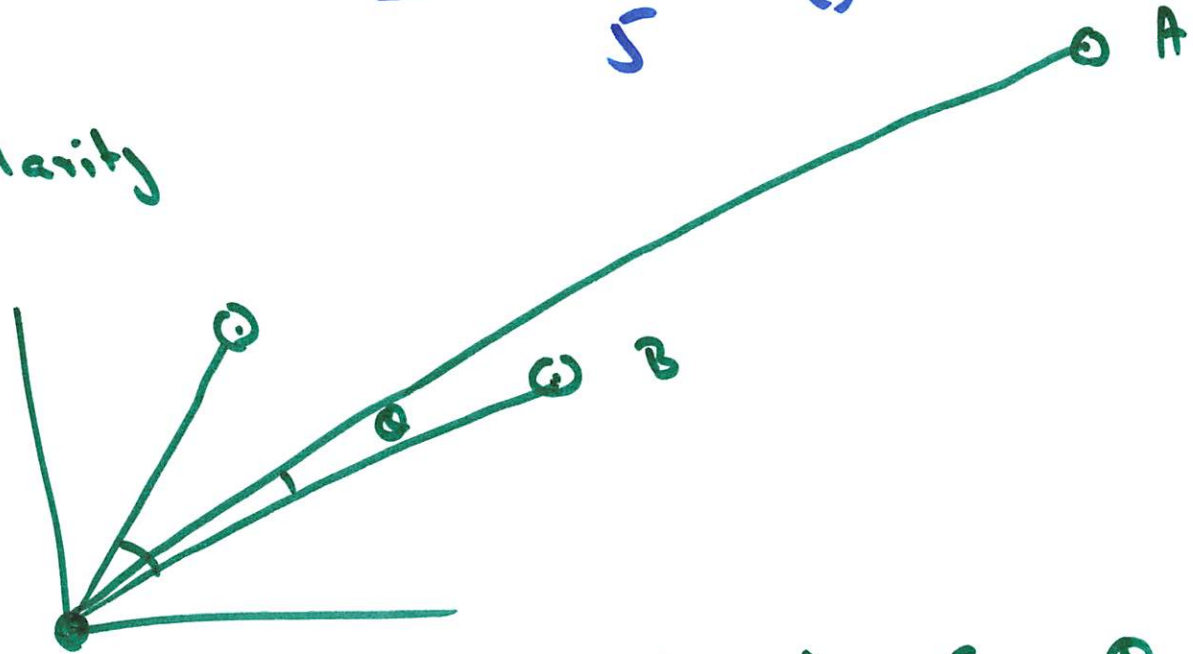


200000×200000

Zucledian



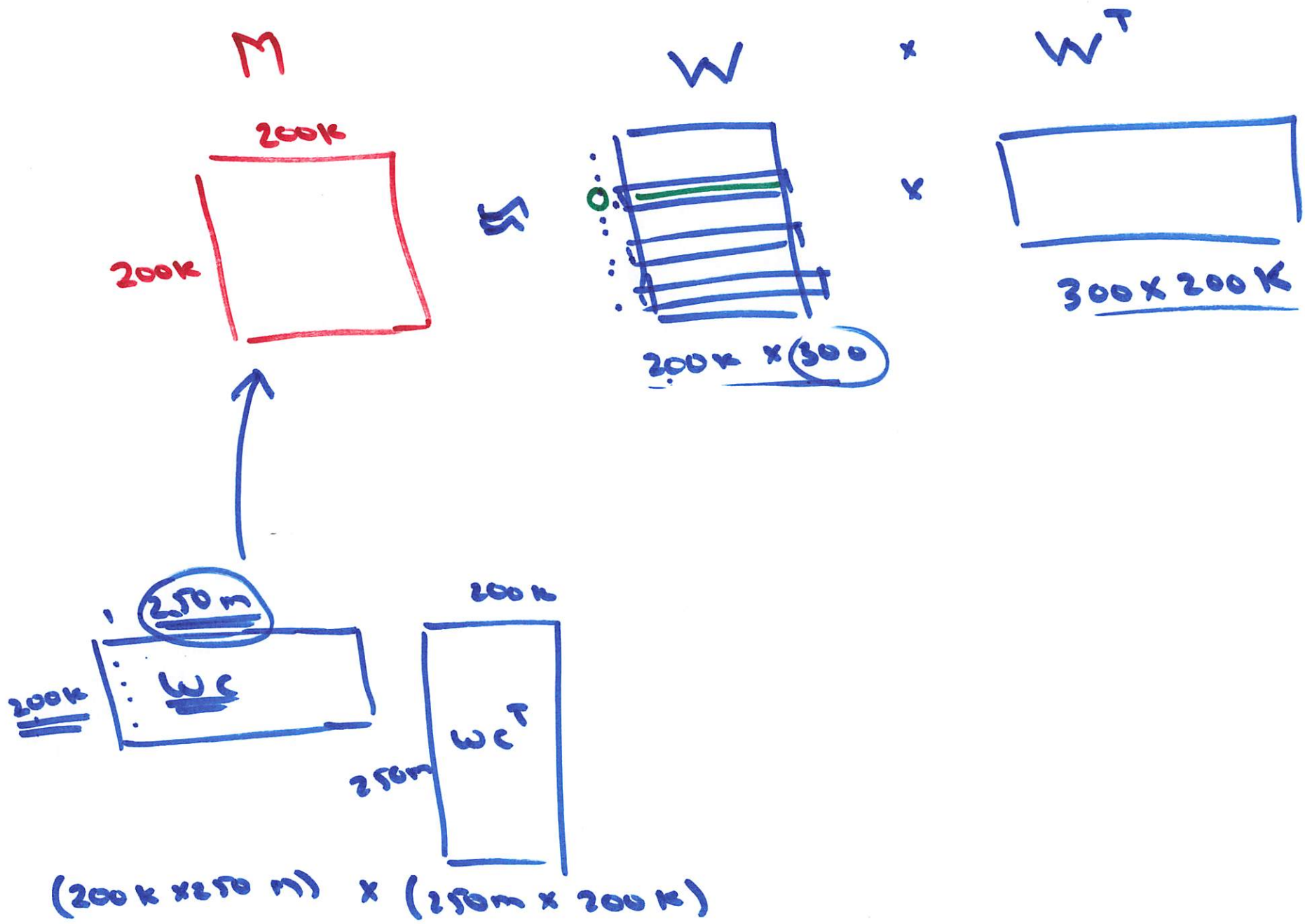
Cosine Similarity



Con θ

$$\text{Similarity}(A, B) = \cos \theta$$

SVD, PCA



~~King : man~~

man : King as woman : ? → Queen

King - Man + Woman = Queen



UK : London as US : ~~NY~~ NY

He : man
King

She : woman
Queen