Model Complexity

word2vec: $O(|C|)$, linear in corpus size

GloVe naïve estimate: $O(|V|^2)$, square of vocab size

But it actually only depends on the number of non-zero entries in **X**. If co-occurrences are modeled via a power law, then we have

$$X_{ij} = \frac{k}{(r_{ij})^{\alpha}}.$$

Modeling $|C|$ under this assumption, the authors eventually arrive at

$$|\mathbf{X}| = \begin{cases} O(|C|) & \text{if } \alpha < 1 \\ O(|C|^{1/\alpha}) & \text{otherwise} \end{cases}$$

where the corpora studied in the paper were well modeled with $\alpha = 1.25$, leading to $O(|C|^{0.8})$. In practice, it's faster than word2vec.