GloVe Objective

$$\mathbf{w}_i^\top \tilde{\mathbf{w}}_j + b_i + \tilde{b}_j = \log X_{ij}$$

Again: two sets of vectors: “target” vectors $\mathbf{w}$ and “context” vectors $\tilde{\mathbf{w}}$.

$X_{ij} = n_{w_i, w_j}$ is the number of times w_j appears in the context of w_i.

Problem: all co-occurrences are weighted equally

$$J = \sum_{i,j=1}^{V} f(X_{ij})(\mathbf{w}_i^\top \tilde{\mathbf{w}}_j + b_i + \tilde{b}_j - \log X_{ij})^2$$

Final objective is just **weighted least-squares**. $f(X_{ij})$ serves as a “dampener”, lessening the weight of the rare co-occurrences.

$$f(x) = \begin{cases} \left(\frac{x}{x_{max}}\right)^\alpha & \text{if } x < x_{max} \\ 1 & \text{otherwise} \end{cases}$$

where α’s default is (you guessed it) 3/4, and x_{max}’s default is 100.