GloVe (global vectors for word representation)

I want a glass of orange juice to go along with my cereal.

c, t

$X_{ij} = \#$ times i appears in context of j.

$\underset{c \quad t}{\underset{\uparrow \quad \uparrow}{X_{ij}}} \qquad \qquad \underset{t}{\underset{\uparrow}{i}} \qquad \qquad \qquad \qquad \underset{c}{\underset{\uparrow}{j}}$

$X_{ij} = X_{ji} \quad \leftarrow$

[Pennington et. al., 2014. GloVe: Global vectors for word representation]   Andrew Ng

Model

$$\text{minimize} \quad \sum_{i=1}^{10,000} \sum_{j=1}^{10,000} f(X_{ij}) \left(\Theta_i^T e_j + b_i + b_j' - \log X_{ij} \right)^2 \longleftarrow$$

$\overset{t}{\downarrow} \quad \overset{c}{\downarrow} \qquad \overset{t}{\downarrow} \quad \overset{c}{\downarrow} \qquad \nearrow 0 ?$

$"\Theta_t^T e_c"$

weighting term

$f(X_{ij}) = 0$ at $X_{ij} = 0$.
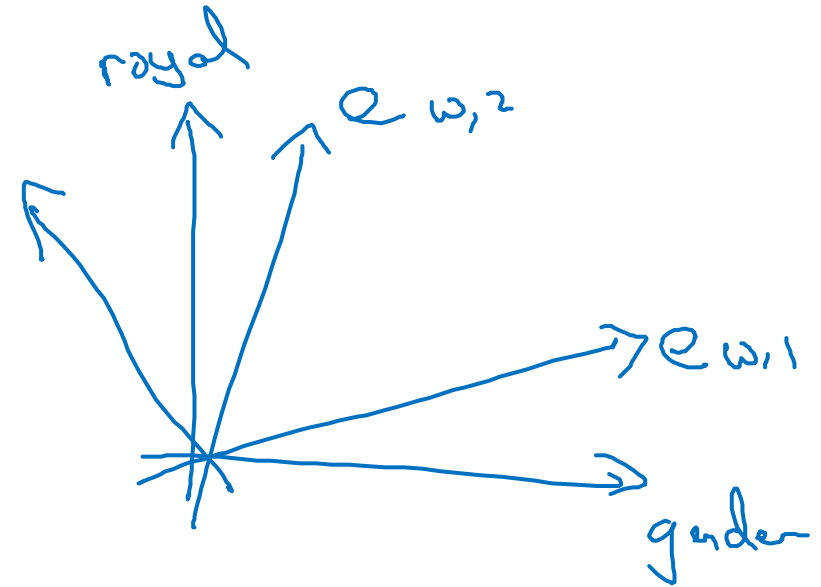
"$0 \log 0$" $= 0$

this, is, of, a, ...

durian

Θ_i, e_j are symmetric

$$e_w^{(final)} = \frac{e_w + \Theta_w}{2}$$

Andrew Ng

A note on the featurization view of word embeddings

	Man (5391)	Woman (9853)	King (4914)	Queen (7157)
Gender	-1	1	-0.95	0.97
Royal	0.01	0.02	0.93	0.95
Age	0.03	0.02	0.70	0.69
Food	0.09	0.01	0.02	0.01

minimize $\sum_{i=1}^{10,000} \sum_{j=1}^{10,000} f(X_{ij})\left(\overbrace{\theta_i^T e_j} + b_i - b_j' - \log X_{ij}\right)^2$

$(A\theta_i)^T(A^{-T}e_j) = \theta_i^T A^T A^{-T} e_j$