Swivel

Swivel Structure

- Factorize PMI matrix.
- Overcome the disadvantage of PMI.
- Also use an objective function $\mathcal{J}$

- Word i, j appear simultaneously

$$\mathcal{J} = \frac{1}{2} f(x_{ij})(U_i \cdot V_j - \mathrm{PMI}(i,j))^2$$

- Word i, j never appear simultaneously(prevent a negative infinity)

$$\mathcal{J} = \log[1 + \exp(U_i \cdot V_j - \mathrm{PMI}^*(i,j))]$$

$$\mathrm{PMI}^*(i,j) = \log \frac{P(i,j)}{P(i) \times P(j)} = \log \frac{A_{ij}/|D|}{A_{i*}/|D| \times A_{*j}/|D|}$$ (A_{word}: frequency of appearance)

$$\mathcal{J} = \log[1 + \exp(U_i \cdot V_j - \log|D| + \log A_{i*} + \log A_{*j})]$$