Likelihood Functions for PLSA vs. LDA

PLSA

$$p_d(w \mid \{\theta_j\}, \{\pi_{d,j}\}) = \sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)$$

$$\log p(d \mid \{\theta_j\}, \{\pi_{d,j}\}) = \sum_{w \in V} c(w,d) \log[\sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)]$$

$$\log p(C \mid \{\theta_j\}, \{\pi_{d,j}\}) = \sum_{d \in C} \log p(d \mid \{\theta_j\}, \{\pi_{d,j}\})$$

LDA

$$p_d(w \mid \{\theta_j\}, \{\pi_{d,j}\}) = \sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)$$

$$\log p(d \mid \vec{\alpha}, \{\theta_j\}) = \int \sum_{w \in V} c(w,d) \log[\sum_{j=1}^{k} \pi_{d,j} p(w \mid \theta_j)] p(\vec{\pi}_d \mid \vec{\alpha}) d\vec{\pi}_d$$

$$\log p(C \mid \vec{\alpha}, \vec{\beta}) = \int \sum_{d \in C} \log p(d \mid \vec{\alpha}, \{\theta_j\}) \prod_{j=1}^{k} p(\theta_j \mid \vec{\beta}) d\theta_1 \dots d\theta_k$$

Core assumption in all topic models

PLSA component

Added by LDA

10