

MCMC for LDA

Lets go full Bayesian

Known: W data

Unknown: Φ latent variables, distribution over words for each topic

Unknown: Z latent variables, topic of each word

Unknown: Θ latent variables, distribution over topics for each document

$$p(\Phi, \Theta, Z|W) \sim \{Gibbs\ Sampling\}$$