Text Mining Homework - Week 3

Aimee Barciauskas, Felix Gutmann, Guglielmo Pelino, Thomas Vicente

May 10, 2016

Exercise 2

(a)

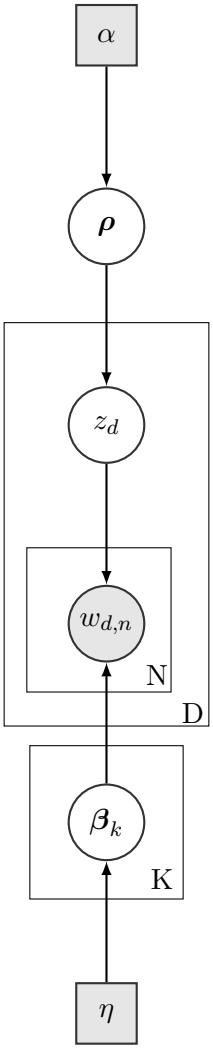

Figure 1: DAG Representation

(b)

The *Markov Blanket* of node V_i in the DAG consists of its parents, its children and the parents of its children. Applying this definition table (1) shows the Markov Blanket for the model.

Node	**Nodes in Markov Blanket**
$w_{d,n}$	$\boldsymbol{\beta}_k$, z_d
z_d	$w_{d,n}$, $\boldsymbol{\beta}_k$
$\boldsymbol{\beta}_k$	$w_{d,n}$, z_d, $\boldsymbol{\beta}_{-k}$

Table 1: Markov Blanket

(c)

1. Choose values for α and η
2. For each s $\in \{1, \ldots, S\}$:

- sample from $\mathbf{P}\left[\boldsymbol{\rho}^s \middle| z_d^{(s-1)}\right] \propto \mathbf{P}\left[z_d^{(s-1)} \middle| \boldsymbol{\rho}^s\right]\mathbf{P}\left[\boldsymbol{\rho}^s\right]$
  $= \prod_k (\rho_k^s)^{\alpha-1}\rho_k^s$
  $\sim \mathbf{Dir}(\alpha + s)$

- sample from $\mathbf{P}\left[\boldsymbol{\beta}_k^s \middle| \mathbf{w}, \boldsymbol{z}^{(s-1)}, \boldsymbol{\beta}_{-k}^{(s-1)}\right] \propto \mathbf{P}\left[\mathbf{w} \middle| \boldsymbol{z}^{(s-1)}, \mathbf{B}\right]\mathbf{P}\left[\boldsymbol{\beta}_k^s\right]$
  $\prod_v \prod_k \beta_{k,v}^{m_{k,v}^{(s-1)}} \prod_v \beta_{k,v}^{\eta-1} \propto \prod_v \beta_{k,v}^{m_{k,v}^{(s-1)}} \prod_v \beta_{k,v}^{\eta-1}$
  $\sim \mathbf{Dir}(\eta + m_{k,1}^{s-1}, \ldots, \eta + m_{k,V}^{s-1})$

- sample from $\mathbf{P}\left[z_d^s = k \middle| \mathbf{w}_d, \mathbf{B}^s, \boldsymbol{\rho}^s\right] = \rho_k^s$,

where $m_{k,v}^{s-1}$ is the number of times topic k allocation variable generates term v in the $(s-1)$-th step.